DOCUMENT RESUME

ED 141 931 EA 009 743

TITLE	Report of Proceeding: Peter Taylor Memorial Seminar on Educational Evaluation.
INSTITUTION	Manitoba Educational Research Council, Ft. Garry.
PUB DATE	77
NOTE	65p.; Not available in hard copy due to light print of original document; Page 45 may be illegible due to small type size
AVAILABLE FROM	Manitoba Educational Research Council, 190 Rupert Avenue, Winnipeg, Manitoba, Canada R3B 0N2
EDRS PRICE	MF-$0.83 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS	Conference Reports; Conflict; *Conflict Resolution; *Educational Assessment; *Evaluation Methods; Models; *Personal Values; *Political Influences; Problem Solving; Program Evaluation; *Social Values; State Departments of Education
IDENTIFIERS	Canada; *Manitoba

ABSTRACT

The bulk of this conference report is made up of three papers. In the first, J. A. Riffel argues that there are two polar orientations to the problem of evaluation in education--the independent empirical and the participatory problem-solving. To date, education has been dominated by the former approach. Only the latter allows for a variety of approaches to evaluation that makes possible a better match between the processes of evaluation and the context in which they are used. In the second, Terry Morrison declares the primacy of values and analyzes the evaluative process, identifying object, criteria, and evaluation framework as components. He sees the effectiveness of value-centered valuation as dependent largely on a more than usual level of tolerance and skill in communication among the participants. Morrison explores the incidence and resolution of conflict within the value-centered framework. The third speaker, Lionel Orlikow, spoke from the vantage point of a member of the provincial department of education and was concerned with three political aspects of evaluation. After discussing three models of evaluation that are possibilities now--external-internal, questionnaire, and standardized tests--he discusses the need for models of consultation and participation. (Author/IRT)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED141931

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

REPORT OF PROCEEDING

PETER TAYLOR MEMORIAL SEMINAR

ON EDUCATIONAL EVALUATION

EA 009 743

1977

THE MANITOBA EDUCATIONAL RESEARCH COUNCIL

$2.00

REPORT OF PROCEEDINGS

PETER TAYLOR MEMORIAL SEMINAR ON EDUCATIONAL EVALUATION

Dr. Peter A. Taylor died suddenly at his home in Ottawa on March 3, 1975. Manitoba educators remember Peter for his contribution to the Faculty of Education of the University of Manitoba and to the Manitoba Educational Research Council. He was a member of the Faculty from 1968 to 1970, and during that time assisted MERC in a number of significant ways. He edited the Manitoba Journal of Education, which under his guidance increased its circulation and began to appeal to a wider audience. With Lil Taylor he wrote the monograph Writing the Objective Classroom Test, and with Tom Maguire he co-authored A Theoretical Evaluation Model, an article first published in the MERC Journal and since widely cited.

Seminar committee:
Colin Walley, chairman
Jack Peach
Odarka Trosky

Lloyd Cavers, staff

Proceedings editors:
Colin Walley, University of Manitoba
Benjy Levin, Manitoba Educational Research Council

Published by the Manitoba Educational Research Council
190 Rupert Ave., Winnipeg, Manitoba

TABLE OF CONTENTS

Introduction	Colin Walley	... i
Evaluation in Education	Anthony Riffel	... 1
What You See is What You Get	Terry Morrison	... 8
Transcripts from Panel Discussion		
	Tom Maguire	...37
	Bob Stake	...41
Comments on Evaluation	Lionel Orlikow	...46
Concluding Remarks	Tom Maguire	...54

INTRODUCTION

The educator has been under siege in the '70's. On the one hand he is accused of usurping too much of the nation's wealth in the vain pursuit of an educational program that would bring happiness for all; on the other he is called to account for a failure to demonstrate that graduates of schooling are competent to read, spell, and calculate.

Under the circumstances it would be understandable for the educator to embark on a program of appeasement; reacting to felt pressure without thought for the eventual consequences. The Peter A. Taylor Memorial Seminar on Education Evaluation was convened so that former colleagues and others who shared a concern for the imperatives of schooling might pause and reflect on them.

The seminar convened at the University of Manitoba, November, 1975, under the sponsorship of the Manitoba Educational Research Council and the Faculty of Education of the University.

Presentors at the Seminar were T. R. Morrison, Head, Educational Foundations, J. A. Riffel, Head, Department of Educational Administration, both of the Faculty of Education, and Lionel Orlikow, Deputy Minister, Department of Education, Manitoba. Revisions of papers which they prepared for the Seminar are included in the text of this document.

Guest contributors, with one exception former associates of Peter's, were Grant Harman, Fellow, Education Research Unit, (Social Sciences), Australian National University; Thomas O. Maguire, Professor, Educational Research Services, Faculty of Education, University of Alberta; Douglas McKie, Head of Graduate Studies, Faculty of Education, University of British Columbia; and Robert Stake, Director, Center for Instructional Research and Curriculum Evaluation, and Professor of Educational Psychology, University of Illinois, Urbana-Champaign.

The theme for the Seminar—"Let's Try and Talk About What's Worth Doing" had a dual focus that is reflected in the papers prepared by Riffel and Morrison.

Riffel, in a paper presented at the opening session—"Evaluation in Education: Toward an Alternative For Small-Scale Studies," raised the need, by implication, for on-going dialogue among those concerned with a particular evaluation; the "Let's Try and Talk" focus of the Seminar.

He argued that most evaluation strategies are based on models in which evaluation is seen as "fundamentally a special form of research." The process "...is to be guided by the rules of empirical inquiry and its outcomes are to be judged according to some strict scientific epistemology." Although he recognized the merits of such an approach to large scale studies he questioned the practicality of the "independent-empirical orientation to evaluation," for the many small scale evaluation studies that are the preoccupation of

practitioners. Rather, he recommended the "participatory problem solving orientation," in part because practitioners are part of the process being evaluated, seek an advocate for their position, are prepared to contend with values, accept their lack of expertise, agree to "muddling through", and assess results in terms of usefulness to themselves in their situation.

Morrison, in a paper that was distributed at the Seminar—"What You See Is What You Get: An Agenda For The Reform of Program Evaluation," stressed the values component in evaluation, the "What's Worth Doing" focus of the Seminar.

He declared the primacy of values and then proceeded to analyze the evaluative process, identifying object, criteria, and evaluation framework as components. He saw the effectiveness of value-centered evaluation as depending largely on a more than usual level of tolerance and skill in communication among the participants. That such must be the case became evident as Morrison explored the incidence and resolution of conflict within the value-centered framework.

There are rewards promised to those who are prepared to accept the likelihood of conflict and deal with it positively: a "prevention of stagnation", having "a medium through which problems can be aired and solutions sought", having opportunity for personal development as well as a source for "personal and group identity", and experiencing "revitalization of norms", "readjustment of the balance of power", and "organizational integration."

But there are difficulties in a value-centered approach to evaluation as Morrison acknowledged. If there is to be "further refinement" of the alternative strategy he proposed, "one that is value based in nature, developmental in purpose and experiential in its process", then "a serious attempt at experimentation and training will be required."

Two themes, a participatory, problem-solving orientation to evaluation, and the relevance of a values-centered evaluation, guided deliberations at the Seminar. The guest contributors—Harman, Maguire, McKie and Stake—commented in turn on the themes. Edited typescripts of the Maguire and Stake remarks are included in the "Proceedings" as representative of the views of the contributors.

The questions on the typescripts to which Maguire and Stake responded came from the floor of the Seminar.

There was strong support from the guest contributors for a participatory, problem-solving orientation to evaluation. The support was largely expressed in Stake and Maguire's repeated reference to the proper demise of judgment rendering evaluators "the superevaluator or Captain Summative" (Maguire). In contrast they supported the idea of an evaluator as a "decision-facilitator" someone with expertise in the structuring, data-gathering, and presentation phases of the evaluation exercises; someone open to a wider range of information about the situation than are the participants, and who has an "obligation to find out what intrinsic valuing there is in the situation." The evaluator "provides a research scene for people who ordinarily are responsible for a program to understand that program better. He's an arranger of things. Sure, he stimulates a little bit, but the decision—the action process—is really quite outside his territory" (Stake).

On the question of values-centered evaluation the guest panel were direct in their view that although values and valuing inhere in any evaluation, it is a poor choice of strategy to attempt to make them explicit.

"It should not be concluded automatically that because values are a central concern to evaluation they should be explicated, studied, dealt with in some sort of resolution form. To set up a situation where there is an examining of the values explicitly involves a constructed as opposed to an intrinsic set of values, and I'm not at all enthusiastic about it." (Stake)

But what of the imperatives for evaluation of education in the everyday world? Some of the answers were to be found in Lionel Orlikow's presentation to the Seminar. In the long term he saw a continuing need "for various types of structures that will promote opportunities for all interested parties—pupils, teachers, parents, and so on—to have a share. If we could develop certain types of models of consultation and participation that would permit discussions, that would permit confrontations, but at least have decision-making taken to the action level—the classroom—it would be possible for information and decisions that are to be made to be translated quickly."

For the moment he saw a much lesser place for participation and values consideration in the evaluation of education.

"Right now I would repeat that the prevailing mood in dealing with the swing of the pendulum will be to reinforce centralism, to reinforce specialization by bringing in more experts, and thereby to exclude or to make more insignificant the role of the individual classroom teacher and the neighborhood in which that teacher works." Orlikow saw a lack of public confidence in education; he saw reaction overriding reflection and in place of qualitative assessment a brewing battle of numbers. "The quantitative approach, for example the use of collective numbers, that will put this school .3 of a norm grade ahead of another will become all-important, although I maintain it would have very little beneficial impact on classroom performance."

The seminar was brought to a close by Tom Maguire. There were those who went away disappointed at having received few answers to the problems they would once more have to face. There were others, including Maguire who in spite of a "sense of confusion, occasional bewilderment", the very thing that may well typify all efforts to evaluate, went away with a "sense of questioning, a sense of exhilaration at having spent the time threshing away at this."

And how else is the harvest to be secured?

Colin S. Walley, Chairman
of the Seminar
January, 1977

EVALUATION IN EDUCATION:

TOWARD AN ALTERNATIVE FOR SMALL-SCALE STUDIES

J. A. Riffel

University of Manitoba

Current thinking about evaluation in education is dominated by empirical concepts and by models which have been devised for use in large-scale, external evaluation studies. Such concepts and models have their place, but their usefulness has been over-estimated. Their usefulness is limited especially in the case of small-scale evaluation studies where educational practitioners such as teachers and administrators are involved in the evaluation process and have a vital interest in its results. The purpose of this paper is to examine some of the limitations on the use of current concepts and models in small-scale studies, and to suggest an alternative conception of evaluation which might be useable in them.

Independent - Empirical Evaluation: The Current Paradigm

Theorists of evaluation in education have patterned their work and their models on the requirements of strict science. They have sought an approach to evaluation which: is objective; separates "facts" from "values"; places a premium on "hard" rather than "soft" data; seeks replicability and generalizability; and requires independence. Viewed in these terms, evaluation is fundamentally a special form of research. The process is to be guided by the rules of empirical inquiry, and its outcomes are to be judged according to some strict scientific epistlemology.

The primary characteristics of this paradigm in evaluation are that:

1. the evaluator is a disinterested, independent observer of that which is being evaluated;

2. with respect to values, and the issues at stake, the evaluator is neutral or non-partisan;

3. some reasonable position on values to guide the evaluation can be obtained. The evaluator is merely a conduit for these values, and treats them as data in the evaluation process;

4. evaluation procedures are rule-guided, following some a priori logic, designed to produce knowledge which is replicable or valid in a psychometric or some broader scientific sense;

5. the results of the evaluation are to be assessed according to empirical criteria of reliability and validity.

The Limits of the Current Paradigm

The independent-empirical paradigm for evaluation seems to have been useful in some large-scale, external evaluation studies, though perhaps not so useful as its adherents would have us believe (see D. A. Mackay, "Through a Glass Darkly: The Realities of Program Evaluation," unpublished paper presented at the annual meeting of the Canadian Society for the Study of Education, Edmonton, June, 1975). This paradigm, however, is of less use in small-scale ventures, where educational practitioners are involved in the process, and vitally interested in the results. There are many examples of such small-scale evaluation ventures in education, for example: evaluating the bilingual education program of a particular school; assessing the educational programs in a school district's open area schools; or appraising the community education program of a particular school. Such ventures are usually undertaken at the request of an educational authority which wishes to know whether some program or activity merits further funding or some extended application throughout the schools in its jurisdiction. Needless to say, the people whose work is being evaluated have a strong interest in the process and outcomes of the evaluation.

Small-scale ventures in evaluation such as the above, which usually involve practitioners in the evaluation process, have a number of important characteristics:

1. the practitioners are immersed in the social and political realities of their situation. They seek the active, first-hand involvement of the evaluator with their work.

2. the practitioners are more likely to seek an advocate for their own basic position than someone who is non-partisan.

3. values are problemmatic, and value is as likely as value consensus.

4. because the practitioners are not trained evaluators, and because of the primary importance of local political realities, the evaluation process has many elements of "muddling through."

5. the results of the evaluation are assessed in terms of their usefulness to the practitioners in the situation.

These characteristics of small scale evaluation studies stand in marked contrast to those of the independent-empirical orientation to evaluation. This contrast suggests that the concepts and models of independent-empirical evaluation are not likely to work in the context of small scale evaluation studies. Needed are additional models which might be useable in the latter situation.

Evaluation Studies: Some Central Issues

The previously mentioned contrast between the independent-empirical orientation to evaluation and characteristics of small-scale evaluation studies raises at least three fundamental question about evaluation in education:

1. what is good evaluation in education?

2. what interests and which people should evaluation serve?

3. what should be the relationship between evaluation and education?

These are normative questions, and they invite a prescriptive response. Evaluators, especially those trained in the psychometric tradition, and educational practitioners such as teachers and administrators, have tended to give different answers to them. These differences are the subject of the followingg analysis.

A Tale of Evaluation in Education! The story is told of an evaluator who had worked for two years appraising the program and procedures of an open area elementary school. He had formulated a number of recommendations and had been urging the staff of the school, somewhat fruitlessly, to implement them.

After months of failure, the evaluator learned he had a terminal disease.

"But I have so much left to do," he told his physician.

"I am sorry," the doctor replied.

The evaluator passed away and regained consciousness in an ethereal and strangely familiar place. He thought it to be the most interesting place he had ever seen. He realized that this perfection was indeed the materialization of his recommendations for his earthly open area school.

As he walked through this Utopia, he say an imposing figure of obvious authority. Approaching the figure, he said, "Sir, I had no idea that perfect plans could reach such a high level of implementation here in heaven."

Somewhat puzzled and angry, the looming figure spat back, "That is not where you are!"

What is "Good" Evaluation? How, then, does one determine what is good evaluation and what is not?

Until recently theorists of evaluation in education have sought their answers to this question in "science". That is, they have patterned their ways of gaining knowledge and making value judgements on models derived from the behavioral sciences. The basic paradigms of science became accepted as the ideal in evaluation—objectivity, "hard" rather than "soft" data, replicability, generalizability and so on. These paradigms, for a while, were accepted as the characteristics of a "good" evaluation. In small scale evaluation projects, however, practitioners usually have a different criterion of "goodness"—and that is usefulness in a particular situation. Might the evaluation lead to improvements in practice or to defensible changes in established ways of doing things? Does it lead to new insights and helpful ways of understanding the situation? When usefulness in a particular situation becomes the major criterion for determining "What is a good evaluation?" the notions of strict science take on less significance. Subjectivity and "soft" data assume major importance. The concerns for generalizability and replicability give way to the need for intimate, first-hand knowledge of the situation being evaluated.

What Interests and Which People Should Evaluation Serve?

Many of the procedures suggested by evaluation theorists represent attempts to disentangle themselves from political pressures and social influences which might influence their work. This search for independence which has its origin in science, had political roots as well. Too many have had their hard, good work subverted by narrow political or self-interest.

Practitioners, for their part, view this search for independence as something of an avoidance strategy. They cannot divorce themselves from social and political pressures—such are the realities with which they must cope daily. They seek the active, first hand involvement of evaluators with their work. Only then, they argue, will evaluators have the intimate understanding of a situation required to make usable recommendations for it. They argue that evaluators have a responsibility to them, not in the sense of safeguarding their self-interests, but in the sense of helping them to improve their work.

What Should Be the Relationship Between Evaluation and Action?

According to evaluation theorists, evaluative research can supply a fundamental base of knowledge and judgements to guide the renewal of educational programs. The potential contribution of evaluation to planned change in education, they argue, is great. Practitioners accept this basic position, but with an important qualification. They recognize that the local political process will overrule the outcomes of an evaluation based solely upon scientific principles. Knowledge is translated into action through social and political processes. Knowing and evaluating are for the sake of doing, and doing is made possible through political action.

The Issues in Perspective

The import of these issues, and the divergent perspectives on them, is that new conceptions of evaluation are required for small scale ventures in educational evaluation. These conceptions, and their basic differences to the traditional pattern, are outlined in the following table:

	Independent Empirical Orientation	Participatory Problem-Solving Orientation
Role of Evaluator	"Disinterested" independent observer of system	Open-minded participant-observer of system
Political Posture of Evaluator	Neutral; non-partisan	Activist; informed partisan
Role of Values	Values as "data"; primacy of descriptive information	Importance of values
Epistemological Rules and Criteria for Valid Knowledge	Replicability or reproducibility	Face-to-faceness as basis of "knowing;" understanding
Procedures	"Rule-Guided"	"Muddling-through"
Criteria for Assessing Evaluation	Rules of strict "sciencing"	Usefulness in illuminating and improving practice

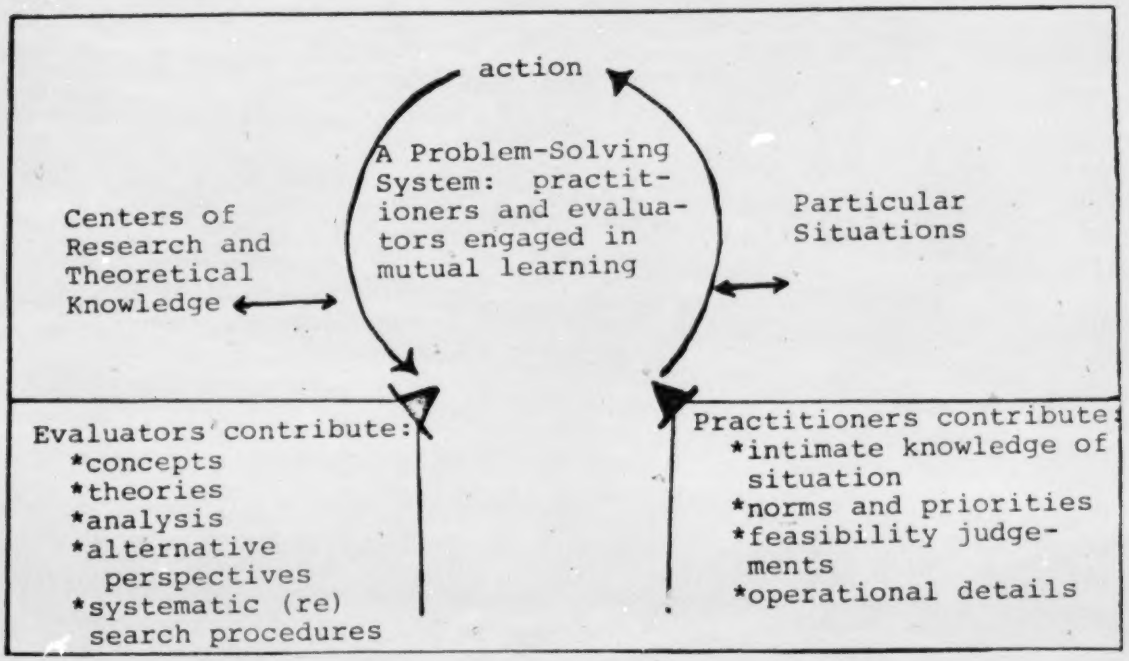
Toward an Alternative: The Participatory Problem Solving Approach to Evaluation

If the foregoing analysis is essentially correct, though somewhat overdrawn, there is need for a set of concepts and models for evaluation that will be useful in small-scale ventures in which practitioners are involved in the process. Such concepts and models are likely to be open pitfalls: there have been too many evaluators who have had to work as "hired guns in local range wars" or who have had their good, hard work subverted by narrow political—or self-interest. These pitfalls, though, should not deter us from the attempt to develop the required models and concepts. We must not confuse impartiality and open-mindedness with aloofness from the social and political realities of a particular situation. Impartiality and involvement are both required.

This hard-to-achieve combination of impartiality and involvement is a characteristic of several established approaches to science—hermeneutics (Habermas, Knowledge and Human Interests, Boston, Beacon Press, 1971) and participant-dialogue techniques (Glaser and Strauss, The Discovery of Grounded Theory, Chicago, Aldine Press, 1967), to name but two. These approaches have a number of characteristics which separate them from strict science. Their implications for evaluation, in contrast to those of strict science, are outlined below:

1. <u>A major portion of future efforts in educational evaluation should be directed to developing a participatory problem-solving orientation to evaluation</u>. At present, evaluation in education is in a state of malaise. The signs of this malaise are many—widespread apprehension and concern about evaluation in schools, the lack of impact which evaluative studies have had on educational practice, and so on. It has come to this state because of over-reliance on, and misapplications of, the empirical orientation to evaluation.

   Needed, then, are additional modes of evaluation, modes which will counter-balance and serve as alternatives in certain contexts to the independent empirical orientation. Such modes would provide other ways of understanding and evaluating education. They would bring evaluation theorists and practicing educators together in joint problem-solving—a rather uncommon occurrence at present. This coming together of theorists and practitioners can be illustrated as follows:



2. <u>North American preconceptions about science should be replaced with the view that there are many forms of scientific activity</u>. Science has been, and will continue to be, an extra-ordinarily creative and valuable human activity. It takes many forms. Habermas, for example, has identified and drawn distinctions among three forms: "strict science," "hermeneutic science," and "critical science". Strict science aims at information which is certain and technically controlled—examples include physics and behavioral psychology. Hermeneutic science is aimed at understanding (<u>verstehen</u>) rather than at

technically controlled information—for example, cultural anthropology. And critical science suggests that information and understanding are not enough—its fundamental aim is the creative use of that information and understanding for the renewal of individuals and their institutions. Its aim is human emancipation through praxis.

All of these forms of science have a place in educational evaluation. What Habermas called "critical science" and "hermeneutic science" are consistent with the participatory problem solving orientation to evaluation. They might be used in a variety of educational contexts—evaluation of a school's program by the staff of the school, a review of organizational procedures within a Faculty of Education, or an assessment of the educational needs of an Indian community. These are essentially local situations, where evaluation is seen as a part of local problem-solving. They are situations in which one finds intense personal involvement, immediate consequences for educational practice, initiation by insiders rather than outsiders, and so on. Here evaluators are blended into a team of problem solvers. The work has the quality of a professional dialogue in joint problem solving between evaluators and local personnel.

Modes of evaluative activity patterned after Habermas' notion of strict science (what I have called the independent-empirical orientation to evaluation) also have their place in certain situations. Take, for example, the need to undertake a province-wide assessment of educational achievement or the need to make rural-urban comparisons of participation in post-secondary education, both for the purposes of reviewing provincial educational policies. In such situations, generalizations based on reliable and valid aggregate data are essential to informed policy making at the provincial level.

There is, then, a need for a better matching between approaches to evaluation and the settings in which evaluation will be conducted. Notions of strict science will be useful in some situations, but not in others. Similarly, notions of critical science will be appropriate in certain contexts.

Summary

Let me summarize my argument. In this paper I have argued that there are two polar orientations to the problem of evaluation in education—the independent empirical orientation and the participatory problem solving orientation. To date, the field of educational evaluation has been dominated by the first. Future work in the field should be directed toward the second. Only then will educators have available to them a variety of approaches to evaluation. This variety will make possible a better matching between the processes of evaluation and the contexts in which they are used.

Notes

[1]Adapted from A. J. Catanese. Planners and Local Politics: Impossible Dreams. Beverly Hills: Sage, 1974.

[2]Adapted from J. Friedman. Retracking America. New York: Anchor, 1973.

[3]J. Habermas. Knowledge and Human Interests. Boston: Beacon Press, 1971.

WHAT YOU SEE IS WHAT YOU GET

An Agenda For the Reform of Program Evaluation

Not so long ago the lack of systematic evaluation of education was seen to be the "problem" of pressing concern. At the present juncture, attempts to systematize evaluation have themselves become a problem for education. A generation of eager and resourceful educators, describing themselves as evaluators, have sought ingenious ways to fashion the educational world into the image of a zoology or psychology laboratory. With such an image ingrained in their minds, these "new breed" evaluators avail themselves to those who would like to know how well various types of people perform in the educational lab on various assigned tasks. The process is, of course, quite appealing to the "scientific" and "technical" mind of today, for it purports to be clean, controlled, dispassionate and uncontaminated by such messy things as values, beliefs, or cultural differences. It is the argument of this paper, however, that evaluation is centrally concerned with just those messy factors. Moreover, its methodology must be found, not in the laboratory of the physical scientist or engineer, but in the political and cultural dimensions of day-to-day relationships between people.

The Concept of Evaluation

Whenever evaluation occurs in any of its diverse forms—comparative assessment, discrepancy analysis, systems functioning or subjective judgement—values must enter in. It is true that when somebody grades apples, he (she) may never make overt reference to any values. But if the procedure were not guided by unspoken values, such as palatability or nourishability, we would be dealing with classification or measurement, not grading or evaluation.

Evaluation can be described generally as bringing four things together in some way:

(i) The object (person, process, idea, product) to be evaluated.

(ii) A valuation, providing a framework in terms of which evaluations can be made.
The framework speaks to the process of conducting an evaluation:
i.e. who is to be involved; what is considered to be evidence and how it is to be assessed; where the evaluation is to be done.

(iii) A criterion of evaluation that contains the standards in terms of which the standing of the object within the valuation framework is is determined.
The criterion mediates the bringing together of the object and the framework. In other words, a criterion allows one to move from establishing that an object possesses certain characteristics to its having certain valued characteristics.

(iv) The situation of the evaluation. This includes the physical setting, the actors, the sense of purpose governing the gathering, and types of things done.

The Prevalence of Values in Evaluation

Since evaluation is a judgement of worth, including within it as a necessary but not sufficient element descriptions of fact, it is impossible to approach either of the above four elements of the process in a totally descriptive manner. The differences between evaluative and descriptive questions in relation to three of the components are illustrated in the accompanying chart.

COMPONENTS OF EVALUATION PROCESS	DESCRIPTIVE QUESTION	EVALUATIVE QUESTION
1. Object of Evaluation.	What are the existing (ideas, products, processes, people) to be evaluated?	What (ideas, processes, products, people) ought to be evaluated and why?
2. Criteria.	What standards already exist in the situation through which to evaluate the (products, processes, ideas, people?)	What standards should be used or developed against which to validly evaluate the (ideas, products, people, processes) and why?
3. Valuation Framework.	Who is involved in the evaluation? What methods will best evaluate existing objects in relation to accepted criteria? Where will the evaluation be conducted?	In relation to the preferred standards, who ought to be involved in the evaluation, where should it be conducted, and what procedures ought to be used?

A process of evaluation, as the chart illustrates, requires that one address value questions to all components of the process. This argument derives not only from the logic of the concept of evaluation itself, but also rests on the fact that values intervene either directly or indirectly in all components of the activity. The next portion of this paper highlights problems associated with value conflict and sketches, as well, a framework within which to approach it.

I Criteria

Education is a valuative enterprise in that some conception of worth is built into whatever is transmitted through the process.* When we talk of facilitating children's learning, we do not mean learning just anything. Helping adolescents learn how to highjack airplanes, for example, would be an unacceptable activity to most people. One basis upon which to determine criteria in evaluation, it follows, is that the standard arrived at be of educational value.** Simply because something exists within an institution such as the school, or is developed through a "curriculum project" with an appealing title, does not, it follows, establish its educational value.

Any attempt to establish the educational value of something will lead one inevitable to an awareness that such an evaluation is a process of social valuing. In theory, as in practice, it involves one or more groups of people assigning values to activities, goals and procedures initiated by others, such as students. Moreover, it contains a particular conception of the types of values to be placed on these activities, goals and procedures. Evaluation, therefore, implies that one must choose from among a range of possible value systems that might give meaning to educational activity. For example, a particular focus of an evaluation can be valued for its efficiency (its ability to get a student from point A to point B rapidly and at low cost), or for its aesthetic qualities (the extent to which it is an experience of beauty and form), or for its contribution to human development (the extent to which it induces qualitative change toward greater differentiation and integration in the person), or for its political effect (the extent to which it increases the power of individuals or groups to make determinations about their own present and future actions). It is not naturally preordained, in other words, that any activity which is labelled as educational be valued in and of itself. To establish such worth, one must refer to a set of extrinsic value premises. The process of deciding upon the criteria (that is, what value or value system) to be used as standards against which to judge the worth of an object of evaluation requires a judgement of extrinsic value. There is no escaping such a value judgement, for even those approaches which use criteria operative in schools as a basis for judgements rest on the unexamined assumption that these standards ought to be valued.

II The Object

Conventional objects of evaluation—certain types of performance, forms of knowledge, dispositions, character traits or propensities—are not necessarily good in and of themselves. They are made valuable, in fact, through the application of certain social constructs. That is, the guiding principles of evaluation—

* See R. S. Peters, Ethics and Education (London, Routledge and Kegan Paul, 1969)

** See B. S. Crittenden, "Educational Value", Journal of Value Inquiry, 1972.

concepts of achievement or skill for example—are cultural constructs which are applied to people. More often than not they represent social rules and cultural norms used by specific groups to define, for themselves and others, expected and acceptable standards of conduct. These conceptions bear remarkable similarity to the concept of "deviance", in that just as other people must define an individual's behavior as "out of line" or "abnormal" for it to be deviant, so too people other than those being evaluated must participate in the definition of whether a particular educational activity is good or bad.

In his discussion of deviance, Howard Becker further clarifies the role of labelling in identifying and evaluating behavior:

> Social groups create deviance by making rules whose infraction constitutes deviance, and by applying these rules to particular people and labelling them as outsiders. From this point of view deviance is not a quality of the act the person commits, but rather a consequence of the application by others of rules and sanctions to an offender. The deviant is one to whom that label has successfully been applied, deviant behavior is behavior that people so label.*

The consequences of Beckers argument for evaluation in education are profound. They imply that a complete understanding of evaluation requires an analysis into not only the groups of people and programs being evaluated, but also the social constructs and cultural norms that make certain things important and other things unimportant. Accordingly, one must evaluate the school or curriculum as an institution that embodies these social constructs and norms as well as evaluating the recipients, the students. Indeed, one must also evaluate the worthwhileness and moral status of the existing constructs and norms themselves. Failure to recognize that evaluation is a cultural act has led to such unjust practices as judging the degree to which groups of students measure up to middle class conceptions of performance and deportment. The continued failure of native students within our public schools is a case in point.

III The Valuation Framework

The current vogue in evaluation procedure and methodology represents a form of systems control and regulation. Essentially, this evaluation approach entails the following steps: One defines a program's educational objectives (preferably in measurable terms); specific procedures are developed to move a student from point A to B (that is, from not meeting to meeting the objectives); evaluation occurs both formatively (along the way) and summatively (at thecompletion of the program) and the results are compared to other programs, in terms of the pre and post test condition of the person, or to the discrepancy between goals and performance. This evaluation provides a "feedback" that supposedly helps the system run more smoothly and efficiently.

In essence, such an approach to evaluation views the process of schooling as akin to industrial production.

* H. Becker, The Outsiders. (Free Press, 1963)

As an evaluation model, such an approach serves a politically conservative function, since it can reinforce what, at times, may be invalid or worthless goals. Several other questionable results flow from such a one dimensional or process-product form of reasoning in evaluation:

(a) The Objectification of Individuals

The process-product reasoning of systems evaluation models often leads to a view of people as "things"—manipulable inputs or abstractions. A latent danger in such a view is that one may, in fact, reify the metaphor and begin to treat people as if they were indeed things.

(b) Goal Displacement

Instrumental reasoning often leads to what sociologists refer to as goal displacement, where means actually become converted into goals. A systems model of evaluation, for example, frequently leads to efficiency being established as not only a desired characteristic of means, but also the end itself. One might value behavioural objectives for their supposed ability to foster measurable outcomes (that is, their efficiency as a means to achieving previously selected ends), but the assertion that such formulations are worthwhile as objectives is arguable.

(c) Avoidance of Value Pluralism

The process-product reasoning in current styles of evaluation provides a smoke screen blurring the essential political and valuative nature of the educational process. Evaluation within this mode, rather than drawing embedded value orientations and conflicts to the level of conscious awareness where they can be analysed, justified and assessed, often ignores the existence of value pluralism and substitutes instead a covering ideology of systems maintenance. These taken-for-granted conceptions, however, tend to be quite conservative in reference to existing educational institutions. Those conservative ideological characteristics described by Alvin Gouldner as integral to structural-functional theory (of which the educational brand of systems analysis is a direct product) can, in fact, be found within current approaches to educational evaluation:*

(1) Treating institutions as given and unchangeable as essentials.

(2) Proposing remedies for them so that they work better rather than devising alternatives to them.

(3) Envisaging no future better than the present—that is, the condition that already exists.

* Alvin Gouldner, The Coming Crisis in Western Sociology (New York: Free Press, 1973)

(4) Explicitly or implicitly advising acceptance of or resignation to what exists, rather than struggling against it.

Evaluation, as has been argued elsewhere by this author, requires that value inquiry be initiated on four levels: (1) value identification (2) value clarification (3) value assessment (4) value prescription.* It is insufficient in evaluation, in other words, to merely identify and help others become clear about values. Critical assessment of those values held by people, particularly the ones used on a basis from which to issue policy or prescribe practice, is also essential. The effectiveness of an evaluator in conducting such a value analysis depends on his/her sensitivity to and skill in exposing, conceptualizing, synthesizing, critizing and assessing values in the various ways they are manifest.

Since values are expressions, in part, of human needs and ideals, analytic acuity alone will not suffice as a foundation for value inquiry in evaluation. The ability to explore perceptual variations, flow with expressions of emotion and appreciate the personal investment in human values is of equal importance. In this regard, the accompanying chart attempts to sketch a framework through which to confront and deal with value conflicts and patterns of value orientation embedded in evaluation situations.

Although specific details related to this approach to values cannot be adequately covered in this paper, it is possible to outline a series of general guidelines for the suggested strategy.

I <u>Effectiveness in Value-Centered Evaluation Demands Knowledge of Factual Information and Concepts Relevant to Value and Moral Decisions</u>.

There are three areas in which such knowledge is important:

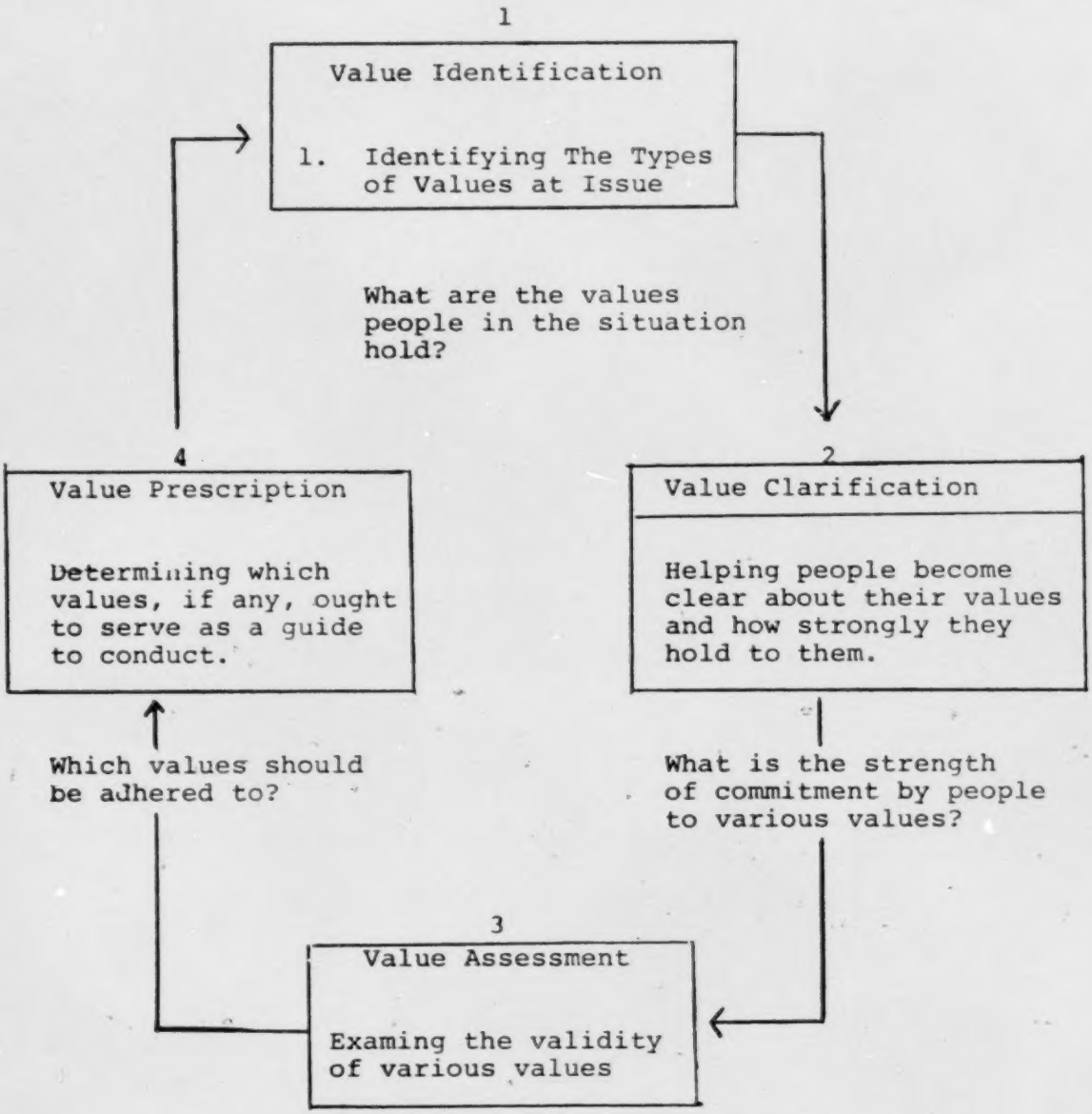
i. the laws, rules, social norms and conventional expectations of society at large and the community groups most directly related to the educational context under evaluation.

ii. Basic data concerning the physical health, safety and well-being of human beings: especially knowledge of those things considered by people to involve risk or threat.

iii. Knowledge of the needs and relative social position of specific groups, especially minority groups and interest groups.

* This approach is detailed in T. R. Morrison and J. A. Riffel, <u>Evaluation and Co-operative Conflict Resolution in Education</u> unpublished manuscript, 1975

A Framework For Value-Centered Evaluation


1
Value Identification
1. Identifying The Types of Values at Issue
What are the values people in the situation hold?
2
Value Clarification
Helping people become clear about their values and how strongly they hold to them.
What is the strength of commitment by people to various values?
3
Value Assessment
Examing the validity of various values
Are those values held by people defensible?
4
Value Prescription
Determining which values, if any, ought to serve as a guide to conduct.
Which values should be adhered to?

II Effectiveness in Value-Centered Evaluation Depends on the Ability of a Participant to Show Concern for Other People as Equals

There are a number of ways of describing this attitude: "regarding other people as equals", "thinking that other people's interests count", "weighing other people's interests equally with one's own", "seeing other people as persons with rights", "taking note of other people's wants, needs," etc.

More specifically, this generalized concern for others requires that the evaluator develop the following orientations:

(1) A concept of the person in which human beings are equal and worthy of respect and due consideration.

(2) A feeling of respect, or benevolence in relation to people.

(3) An action-frame geared to helping others or, at least, avoiding harm to them.

III Effectiveness in Value-Centered Evaluation Relates to the Participant's Ability to be Aware of Feelings and Emotions in Others and in Oneself

This element includes not only awareness of existing emotions, but also the ability to know what emotions a person is likely to experience in future. This capability especially requires that one be able to detect signals of various emotional states.

a) through what a person "says" he/she feels

b) by characteristic symptoms, e.g. sweating

c) through actions patterns

d) by characteristic circumstances, e.g. dangerous situations, contexts of despair.

One of the essential abilities in this area is to have facility in a "language of emotions", that is, to be able to accurately describe others' feelings and one's own through language (e.g. "blue", "jealous","love", "hate","tight", "loose", etc.)

IV Effectiveness In Value-Centered Evaluation Depends on the Ability to Actively Listen and Respond to Others' Concerns and Problems

Evaluation activities in education necessarily involve one in a flow of signals from other people. The ability to sort, classify, and otherwise code these messages, particularly so that the intent of the sender is not swamped by ones own perceptions, is crucial in uncovering the value base of communication.

Evaluators must listen in an active rather than passive mode, however, for their audience expects critical commentary on what is being communicated not just acknowledgement of having received the message. In this regard, the response

style employed by an evaluator is crucial. As an example, consider the following types of response style which are available to an evaluator:

(1) Evaluative

A response that indicates the receiver has made a judgement of relative goodness, appropriateness, effectiveness, or rightness of the sender's problem. The receiver has in some way implied what the sender might or ought to do.

(2) Interpretative

A response that indicates the receiver's intent is to teach, to tell the sender what his problem means, how the sender really feels about the situation. The receiver has either obviously or subtly implied what the person with the problem might or ought to think.

(3) Supportive

A response that indicates the receiver's intent is to reassure, to pacify, to reduce the sender's intensity of feeling. The receiver has in some way implied that the sender need not feel as he does.

(4) Probing

A response that indicates the receiver's intent is to seek further information, provoke further discussion along a certain line, question the sender. The receiver has in some way implied that the sender ought or might profitably develop or discuss a point further.

(5) Understanding

A response that indicates the receiver's intent is to respond only to ask the sender whether the receiver correctly understands what the sender is saying, how the sender feels about the problem, and how the sender sees the problem.

The foregoing response styles in themselves are neither good nor bad. It is the over- or under-use of any type which reduces its utility or the failure to respond appropriately that interfaces with the communication flow. When is each response style most appropriate in evaluation?

In the early phases of evaluation the two response styles to be most sensitive to are the understanding and evaluative. Essentially, the understanding response requires that when an individual sends a message, and that message is paraphrased in fresh words with no alteration in fundamental meaning, the sender will tend to expand upon or further search out new ideas, feelings, and attitudes contained in the messages. Thus both trust and creativity can result.

A major barrier to productive communication and mutual understanding is premature judging or evaluating of the messages of a sender. For this reason, one should avoid over-use of evaluative responses in the initial stages of an evaluation. Evaluative responses, however, are quite appropriate when one is asked to make a specific value comment.

When a person tries to discuss an issue with you, that you do not understand, probing responses are helpful. They are also helpful if you feel that the sender does not see the full implications of his or your statements. Supportive responses are effective when a person needs to feel accepted or needs enough support to engage in problem-solving. Interpretative responses are particularly useful in confronting a person with the potential impact of his behavior on you or others.

V Effectiveness in Value-Centered Evaluation Depends on the Willingness of Participants to Face the Value Issues, and Make Value Judgments in Terms of Their Own and Other People's Interests

In a way, this postulate sums up the model put forward. This is, in a sense, a question of will and integrity, and without it all other efforts are self-defeating. Two factors, however, are important in making this will operative: (i) The avoidance or control of modes of thought which tend to found value judgments on bases other than other people's interests, e.g. obedience to authority, conformity to a social group, non-rational obedience to rules, expedience, impulsiveness. (ii) The commitment to act upon and take responsibility for judgments which prescribe action for oneself and others; to recognize that value choices are indeed your value choices with direct implications for your behaviour.

When one speaks of education, there is presupposed either an ideal of a person toward which the process is leading, or a social ideal for which education prepares one. In either case, underpinning the idea of education is a set of values or norms which provide a sense of purpose for the process. It makes no sense to talk of a system of education—or to evaluate aspects of it for that matter—without addressing oneself, at the same time, to values, personal ideals and social visions. For this reason, evaluation as discussed in this paper, refers to a deliberate judgment of worth.

IV The Evaluation Situation

Evaluation takes place in a social context—in a family, small group, community, institution, or a civilization—that has developed techniques, symbols, categories, rules and values which are relevant to the effectiveness of the evaluation. Hence to understand the events that occur as part of an evaluation process, one must comprehend the interplay of these events within the context in which they occur.

Most educational evaluations are based upon models which guide data collection. These models—and there are many—mostly focus on the "treatment" or program under review, rather than the larger context in which the treatment exists. In addition, they usually adhere to either the canons of experimental design or a form of systems analysis. This dedication to experimentalism and systems, both of which have control of the environment as one of their prime targets, often produces a preformed specification of what is to be studied and how it will be studied. In each case, the problem to be researched is set before data is collected. It is little wonder that such studies produce "predictable" results.

In contrast to the experimental or systems of models of evaluation, one can put forth the idea of evaluation as an experiential learning process; a learning process for both the evaluator and those subject to the evaluation. In this type of evaluation, the research problem itself is a problem in that it is not defined a priori to the collection of data, but results in part from data collection.* The evaluator does not ignore theories, concepts or models, but uses them as searching devices and probes. Evaluation, construed in this way, is geared more to the generation of problems and in-practice solutions rather than the verification of hypothesis or validation of programs.

This type of description can aid the evaluator in decision-making only if all dimensions of a situation are accounted for: actors, behaviour, context, setting. Actors are categories of people, while behaviours are categories of things done. A context can be defined as the "sense of purpose" for the gathering of people, while the setting is the physical environment. These four elements combine to form a situation. An evaluator who seeks to understand accurately the dynamics of the situation under review must describe all of these dimensions. By omitting a dimension in ones description of the situation, the evaluator runs the risk of evaluating his/her perception or construction of the problem rather than the problem as experienced by those in the situation. It is this fault in evaluation methodology which often produces such remarks as "If you really fully understood our program, you wouldn't make such recommendations" by those who are subject to evaluations.

What occurs in an evaluation situation is not just a function of what is in peoples' heads (ie. their personality characteristics, potential for learning, or espoused values etc.). It is a function of where they are (the setting), why they are there (the context) and who is there with them (the actors present). Different things can happen in other words, in the same physical setting depending on the context and actors present.**

The methodological technique which offers most promise for incorporating the contextual complexities of the evaluation process is ethnography. This technique differes from traditional data collection methods employed in evaluation which center, for the most part, on the acquisition of quantifiable evidence derived from a priori psycho-statistical models and utilizes, instead, direct observation to obtain qualitative data.*** The evaluator, as a trained ethnographer, directly observes human activity and interaction in an on-going naturalistic fashion. The recording ranges from no recording during observation and total recall later, to note-taking during observation, or tape recording of events as they are being observed. What is produced from this process is a protocol of what occurred.

In contrast to research methodology drawn from experimental psychology or functionalist sociology, ethnographic modes of inquiry in evaluation focus

* D. B. Everbart, "Problems of Doing Fieldwork in Educational Evaluation", Human Organization, Vo. 34, Summer, 1975, No. 2 pp. 206-207.

**See B. A. Tindall, "Ethnography and the Hidden Curriculum In Sport", Behavioral and Social Science Teacher, Vol. 2, No. 2, 1975.

on the development of hypotheses, rather than the testing of them.* The evaluator as observer begins with a tentative set of components and concepts which serve as a rough guide in organizing data collection. It cannot be over-emphasized that this list of components is structural not judgmental in nature, its raison d'etre being to help one order observation. Below is a sample of such a category system based on a modification of Jules Henry's classic Cross-Cultural Outline of Education.**

---

1) On what do the various processes (educational, political, economic, etc.) operative in the situation focus?
2) How is information transmitted and interpreted?
3) Who educates whom?
4) How does the person(s) being educated participate?
5) How does the educator participate?
6) Are some things experienced by some and not others?
7) Are there discontinuities in the educational process?
8) What forms of control over space, time, behavior and the situation exist?
9) What is the relationship between intent and the actual results of education in the situation?
10) What self-concepts are reinforced?
11) What standards are used to determine normality and deviancy?
12) How long does the educational process last?
13) What basis is used to determine success within the context of the educational process?
14) Are there any assumptions about the need for prior knowledge or experience in order to participate in the educational process?

Regardless of the subject of an evaluation, any set of categories to be used within an ethnographic study should be drawn from at least four sources:

(i) prior knowledge of the context under examination

(ii) initial and on-going observations of the context itself

(iii) theoretical and conceptual orientations of the observer

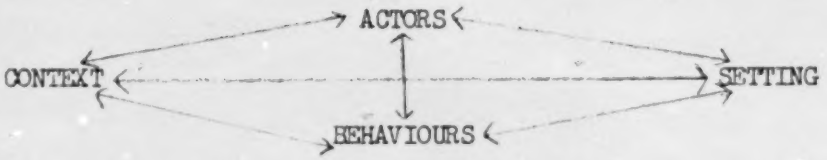
(iv) meaning structures and category systems held and used by people in the context.

As one can see from the foregoing list of sources, the observational categories used in ethnographic inquiry form an evolving and open system which should have as one of its prime traits responsiveness to changes in the context itself.

* See Ray Rist, The Urban School (Yale University, 1974)

**Jules Henry, "A Cross Cultural Outline of Education", Current Anthropology, 1960, Vol. I. pp. 269-305.

Once enough data has been collected to give one a picture of the situation, a more focused analysis can be undertaken. Many forms of analysis exist to aid in this stage. As a proposed approach, however, this writer would suggest a dissection of the socio-cultural system under investigation into its generic parts: actors, behaviours, context and setting. These components can, then be related to one another in various ways so that a set of alternative reconstructions of the system emerge.



The process of reconstruction requires that one focus attention on particular components of the situation while holding others constant, for example, centreing attention on categories of behaviour in particular contexts, while holding actors and the setting constant, or focusing on categories of actors performing types of behaviour, while holding the setting and context constant. Over time, such focused analyses can produce patterns within the evaluative situation. For example: certain types of actors in certain settings and for certain reasons may be accorded preferential treatment by other actors, etc.

The reconstruction of such patterns aids the process of evaluation in at least five ways: (1) It minimizes the tendency to distort evaluative situations to fit the pre-conceived models or hypotheses of the evaluator or evaluative design. (2) It helps one move closer toward the goal of evaluating the situation as it is actually being experienced by those within it. (3) It broadens the information base from which, in part, the necessary normative judgments inherent in evaluation can be formulated and issued. (4) It increases the potentiality of evaluation, as a process, becoming a learning and problem-solving experience for those involved with or affected by it. (5) It increases the probability that recommendations for change, modification or redevelopment will be formulated with greater sensitivity to the constraints, processes and nuances within the situation to which they are addressed. To this extent, at least, the possibility of acceptance of such recommendations increases.

Evaluation and Cultural Conflict

When evaluation is approached as a value-based activity rooted in social interaction, it becomes readily apparent that students, teachers, parents and various members of the general public have differing, and sometimes competing, categories around which they want data placed and judgments framed. These categories reflect different assumptions and value premises about the qualities of good schooling.* Thus, some people want schools to stress the three R's; others want schools to center on inter-personal and human relations. Accordingly, each group feels that schools are to be evaluated in terms of the degree to which they perform effectively in their preferred categories. Changes in schools, it follows, are often a product of pressures generated by shifts in the criteria or value bases upon which parents, teachers and students evaluate schooling. This shifting and competing nature of evaluative categories inevitably forces the evaluator to face squarely several value questions: How does one choose between competing value premises in evaluating public education? How can one adjudicate among competing conceptions of good schooling and, indeed, their underlying conceptions of the good life? How does one respond to change within evaluative categories?

*L. Fein, The Ecology of Public Schools. (Random House, 1972).

The first step in answering questions such as those above is to recognize that the role of evaluation is not so much to dampen as to bring to life variations in cultural and professional orientations toward education. Specifically, this means that evaluation activities should be designed and conducted so as to: (1) help people clarify, analyse and justify their value orientations in regards to education (2) make available to interested groups a range of alternative choices and preferences (3) contribute to the establishment of those conditions through which value orientations can be expressed in action (4) foster those values that underpin cultural pluralism, yet do not commit oneself to blind relativism - freedom of choice, equality, respect for differences, due process.*

In pursuing these goals, evaluation activities will necessarily involve one in value analysis and decision-making on at least three levels:

1) Individual values: This level of analysis requires that the evaluator focus on those values which individuals hold and have a bearing on the object of the evaluation. A parent's view of authority relationships in the home, for example, may affect his/her willingness to accept the norms of conduct implied in "open education."

2) Group Values: This level of analysis centers on the values subscribed to by various groups connected with an activity under evaluation. So-called academic departments within schools, for instance, often espouse particular value orientations, eg. the view that the most valued knowledge is contained within disciplines. Informal groups and cliques within schools also reflect particular values. Analysis of group values, when juxtaposed against the values of individual members, often expose areas of incongruity and tension. Consider the case of a teacher who claims to want more freedom for students but who, as a member of a department, acquiesces to the group's ethos of control.

3) Societal Values: This level of analysis concerns society as a whole. It extends beyond specific cultural groups to relations between individuals and society, and between cultural groups. An evaluation of second language instruction in Canada today cannot be complete unless one addresses such policies as bilingualism, biculturalism and multi-culturalism.** Analysing values at the societal level will also force one to dissect the stated and assumed relationships between school and society. The particular perception of that relationship held by various publics, whether it is defined in terms of social reform, cultural compensation, social control or

* J. Klein, "Cultural Pluralism and Moral Education", The Monist. October 1974. pp. 691-693.

**See A. Wolfgang ed., The Education of Immigrants In Canadian Schools, (Toronto: OISE, 1975).

therapy, will affect not only what is attempted in schools, but also how it can be evaluated.

An evaluation framework which recognizes cultural pluralism, and is conducted on all three levels of analysis suggested, will inevitable lay bare conflict and tension in, at least, the following areas:

(a) between individuals and their membership groups

(b) between groups

(c) between individuals and groups, and between groups and society

(d) between stated preferences and actual behaviour.

The fear among educators and others that these patterns of conflict will in fact surface, and lead, it is assumed, to destructive consequences, is one reason for the rather persistent avoidance of the value and cultural foundations of evaluation. This protective and nurturing stance toward educational evaluation reflects a view of conflict as a negative or "pathological" force, something to be controlled, managed or eliminated from the evaluative process. It is the argument of this writer, however, that conflict is both embedded in and essential to effective evaluation and, if approached creatively and co-operatively, can generate numberous constructive consequences.

Evaluation and Co-Operative Conflict Resolution

Conflict exists when incompatible activities occur. The activities may originate at the personal, group or institutional level. Since an essential ingredient in the process of evaluation is conflict, one's view of conflict itself becomes a crucial factor affecting not only <u>what</u> will be evaluated, but <u>how</u> the evaluation will proceed. Consider the following areas in evaluation about which parties are likely to disagree:

(1) what should be the focus of the evaluation?

(2) whether existing goals or objectives are worthwhile?

(3) who should be involved in the evaluation?

(4) how the evaluation should be conducted, eg. what instruments to use, what data to collect, the time and place?

(5) how the results of the evaluation should be used and who should have access to them?

The typical response to conflict in education is avoidance. There would seem to be at least four reasons for this:

1. an unwillingness to acknowledge the legitimacy of other interests.

2. fear of the destructive effects of conflict.

3. lack of awareness of the constructive effects of conflict.

4. hesitancy to consider or undertake fundamental structural changes in public schooling.

At best, these responses help educators keep the lid on potential sources of disruptiveness. At worst, they establish those alienating forces which actually escalate the level of conflict or lead to the immobilization of educational organizations.

Toward a Positive View of Conflict

Undoubtedly, conflict, when approached within a competitive context, can generate certain destructive consequences. When approached co-operatively, however, conflict can stimulate human development and educational innovation. This is particularly the case in the process of program development which rests, for its effectiveness, on a style of thought which links thesis—antithesis and synthesis, and, in which, conflict is an essential mechanism.

Conflict, when generated and guided through a process of evaluation, has the potential to bring many benefits and positive outcomes. Consider, for example, the following:

1. The Prevention of Stagnation through Stimulation of Interest and Curiosity

   Evaluation—as a process of conflict resolution through social interaction—by bringing conflicts to the level of conscious awareness, allows different ideas and viewpoints to mix and, thus, contributes to the creative process of program and organizational development. Continuous evaluation, it follows, is one way of maximizing creative potential.

2. The Provision of a Medium Through Which Problems can be Aired and Solutions Sought

   Continued repression or denial of conflict often leaves a set of underlying problems untouched, a situation which can affect adversely the motivation and level of commitment of persons involved in activities like curriculum development and implementation. Evaluation, by helping parties conceptualize and discuss latent or manifest problems can itself be a first step toward creative problem-solving.

3. A Stimulator of Personal Development

   Intrapsychic conflict, experienced as various forms of dissonance, has been isolated by numerous social scientists as potentially developmental in its impact.* Constructive resolutions of such conflicts lead to increased understanding of and sensitivity to self, others and the physical and social world. The process of evaluation, if approached as a form of constructive conflict resolution, offers the participants an opportunity to compare and contrast their views of the world with the views of those around them. Constructive

*See: W. Rohwer, P. Ammon, P. Cramer; Understanding Intellectual Development, (New York: Dryden, 1974).

resolutions of the resultant intrapsychic conflicts generated in this process can lead, in the case of the teacher, to significant gains in professional awareness and development.*

4. A Source of Personal and Group Identity

Conflict, by demarcating groups from one another, engenders a sense of group loyalty and, thus, offers a context for the formation of personal identity. Properly guided - that is, preventing the development of a form of group ethnocentrism - such an awareness of group identity can bind a staff or teaching team closer in a quest for solutions to common problems. In innovative programs, the establishment of a "community of the committed" provides a counterweight to the frequent emergence among teachers of feelings of cynicism, alientation, and powerlessness.

5. The Revitalization of Norms

Conflict, when guided constructively through an evaluation process, can revitalize existing norms in a group or institution, or contribute to the development of new ones. Viewed in this perspective, conflict provides a mechanism through which to adjust norms so that they are adequate to new environmental conditions. By facilitating the creation or revitalization of norms, such a process can increase organizational flexibility and thus help insure the continuance of the institution under rapidly changing conditions. Rigid systems, by avoiding or suppressing conflict, eliminate a useful early warning signal and, thereby, increase the possibility of either territorial defensiveness or breakdown in the face of social change.

6. Readjustment of the Balance of Power

Conflict, exposed through a process of evaluation, can also serve as a means of ascertaining the relative strengths of antagonistic interests and, in this way, constitues a useful mechanism for the maintenance or continued readjustment of the balance of power. Once the respective power of the groups is identified, a new equilibrium can be sought.

7. Organizational Integration

In open systems, conflict resolution can facilitate a constructive release of tension and thus exercise a stabilizing and integrating function. By permitting immediate and direct expression of rival perceptions, evaluation provides organizations

* See: H. Thelen, "A Cultural Approach to In-Service Education" in L. Rubin's Improving In-Service Education, and T.R. Morrison and K. Osborne, "Program Development; An Example for Teacher Education", Manitoba Journal of Education, Vol. 10, No. 1, pp. 22-39.

with automatic readjustment mechanisms and resources for dealing fairly with sources of dissatisfaction.*

Guidelines for Co-operative Conflict Resolution

The evaluator, whether functioning as an outside expert, in-house observor, or as part of a team, often occupies the social role of mediator. In such a position, he/she frequently faces the difficult problem of how to intervene therapeutically, and for developmental purposes, in situations which tend to be deadlocked or unproductive because of misunderstanding, hostile attitudes, misperceptions or the inability to generate mutually satisfying solutions. The development of creative approaches to educational evaluation, within a world marked by tightening interdependence and an increasingly common experience, requires that educators look beyond the walls of their own specialities and roles to the pool of alternative strategies available in the community at large. In this regard the approaches to conflict employed by such people as family therapists, labour conciliators or film directors offer evaluators a refreshing set of role expectations and techniques regarding the managment of conflict.** Some of the recurring strategies within these field provide the framework for an emerging framework of conflict management of use to educational evaluators. At the same time they suggest a dramatic alteration in the types of knowledge, skills and experience which are essential to a person involved in evaluation. In future, evaluation should be guided less by the canons of experimentalism, quantification and control, and more by the ethical, subjective and qualitative dimensions of social interaction. Evaluators, in other words, must immerse themselves, rather than stand apart from, the cultural environment of education.

Evaluation as Co-operative Conflict Resolution: A Framework for Action

The creative use of conflict, which is either generated or exposed through evaluation, requires that the conflicting parties be made sensitive to the nature of the nature of the conflict at issue as well as those strategies which are available to approach it. The following model presents a framework through which one can describe, diagnose and creatively employ conflict within an evaluation process. An evaluator, when functioning within the parameters of the model, should be guided by several objectives, each of which describes an orientation toward conflict.

1. Helping Parties Identify and Confront the Issues Underpinning Their Disagreement

   Conflicts are often suppressed and avoided, and occasionally displaced onto other issues. Several techniques are available to help in bringing this latent conflict to the surface: being present as a supportive yet prodding third party; helping to alter the structure of power so that all views are given a hearing; identifying and framing key issues; and establishing favourable conditions for confrontation by eliminating a win-lose definition of the situation.

* See R. Smuck and P. Runkels, Handbook of Organizational Development in Schools (NY: National Press, 1974)

** See K. Kressel, Labour Mediation: An Exploratory Study. (NY: Vintage, 1974), J.S. Coleman, Community Conflict (NY: Free Press, 1957), D.E. Walton Interpersonal Peacemaking (NY: Addison-Wesley, 1969)

2. Helping Remove Blocks or Distortions in the Communication Process

Often evaluators report that part of the problem in a setting is that people talk around, but not to each other and the issues of concern. A central role of the evaluator, it follows, is to "free up" the lines of communication. In this regard, several techniques, from organizational development to group decision-making and problem-solving skills, are available. The point remains, however, that evaluators require as much training in human relations as any "helping professional".*

3. Helping to Establish an Agreed Upon and Just Set of Norms to Govern Interaction.

Mere talk will never lead to effective conflict resolution. Once people begin to discuss freely, it becomes crucial for an evaluator to work with them in establishing the rules by which their talk will be collectively understood, assessed, and, in cases, discarded. To establish such conditions requires that the evaluator perform the role of a jurist. He or she must insure that the rules are clear, adhered to by all, seen to be justly derived from group concensus, and are applied to all. As one can imagine, the ability of an evaluator to be both within and outside of a situation as the need arises is a skill of prime importance.**

THE MODEL

I. Identifying the Type of Conflict

Not all conflict or conflict situations are the same. Misinterpretation of the nature of conflict can lead to evaluation being conducted at the symptomatic level of the application of inappropriate techniques. In this regard Morton Deutsch has developed a typology of conflict which has particular relevance to the educational context. Deutsch distinguishes conflicts in terms of the relationship between the objective state of affairs and the state of affairs as seen by conflicting parties.*** Using this scheme, six types of conflict can be characterized, each of which poses special problems for evaluation.

* See. D. Johnston, Group Theory and Group Skills, (New York: Prentice-Hall, 1975).

** See Morton Deutsch, The Nature of Human Conflict, (Yale: University Press, 1974).

***Ibid., pp. 1-20.

1. Veridical Conflict

This conflict exists objectively and is perceived accurately by all parties. It is, in a sense, true conflict. (for example, one teacher wants to use an available room for a student study centre and another wants to convert the same space into a staff planning office) Veridical conflicts are difficult to resolve amicably unless there is sufficient basis for cooperation between the two parties or agreement to accept impartial arbitration of a third party.

2. Contingent Conflict

Here the existence of conflict depends upon readily rearranged circumstances, but this possibility is not recognized by the conflicting parties. If an evaluator can provide or uncover available alternative resources for satisfying the conflicting "needs", then this type of conflict can be resolved. As well the evaluator may have to prevent loss of face for either party. It should be clear that handling veridical conflict requires the evaluator to play a different role than if he/she was facing a contingent conflict.

3. Displaced Conflict

Here the conflicting parties are arguing about the wrong thing. A senior administrator may quarrel with a project director in his school over inefficiency in managment. This may be, however, merely a displacement of an unexpressed conflict over styles of interpersonal relations. The manifest conflict, thus, is a symbolic expression of the underlying conflict. Often manifest conflicts can only be resolved temporarily unless the underlying conflict is dealt with.

4. Misattributed Conflict

In this form the conflict is between the wrong parties and thus usually centres on the wrong issues. The misattribution may be unwitting, as in the case of a principal who blames a teacher for something he instructed that teacher to do, or it may be fostered deliberately by those who can gain from it. "Divide and conquer"is a well known strategy in this sense. A central role of the evaluator in this setting is to reduce missattribution so that effective co-operation can ensue.

5. Latent Conflict

This is a conflict which should be occurring but is not. Displacement, misattribution, fear, denial or repression can all prevent true conflicts from surfacing. Evaluators interested in educational and social change must often attempt to convert latent conflict into conscious conflict. It is important that fair, equitable and reasonable conditions be established through which to deal with the resultant conflict.

6. False Conflict

This is the occurrence of conflict when there is no objective basis for it. Conflict of this sort rests usually on misperception or misunderstanding. A latent danger in this form of conflict is that it will operate as a self-fulfilling prophecy and become the dynamic element leading to an escalation of real conflict. Evaluators, in paying close attention to the type of conflict at hand, must realize that conflict is a complex phenomenon. There may be veridical conflict on one issue and misattributed conflict on another, both occurring in the same context. Yet one thing is clear, evaluators require more than skills in data collection and manipulation to be effective. They need, in this regard, not only acute perceptual and conceptualizing skills but also the willingness and ability to approach situations openly, from a variety of different angles, using a range of different strategies of intervention as the situation demands. Any evaluation model, to be of use in a social context, it follows, must be dynamic not static, and stress learning processes and alternative action strategies.

II Establishing the Source of the Issue

A conflict, whatever form it takes, will usually focus on one of several types of issues. The nature of the issue will, moreover, affect both the form of the conflict as well as the type of evaluative intervention required. Five types of issues frequently encountered in evaluation situations are noted below:

(i) Control Over Resources

In educational settings, such resources as space, money, power, prestige, staff, materials and so forth may be seen to be nonsharable. If two or more parties seek exclusive use of possession of a resource, conflict is likely to emerge. Conflicts such as these are difficult to resolve unless extra resources can be generated or a new basis for sharing is worked out. In some cases, evaluators can help people see that the particular tasks can be effectively accomplished in alternative ways without reliance on the resource at issue.

(ii) Preferences and Nuisances

Conflicts often arise because the activities or tastes of one person or group collide with those of others. A teacher constantly yells and shouts at students in the room next to you. The issue is not whether one has the right to his preferences, but whether, in the exercise of that right, a nuisance is created for you. While segregation or avoidance can be used to deal with such conflict, the latent danger is that it will become a symbol of deep power struggles.

(iii) Beliefs

As the earlier section of this paper indicates, value

conflicts form a central focus of evauative activities. Value conflicts can be over isolated issues or take tha larger form of religious and ideological conflict. As Morton Deutsch schrewdly observes, "it is not the differences in values per se that lead to conflict but rather the claim that one value should dominate or be applied generally, even by those who hold different values."*

(iv) Beliefs

Conflicts often arise over what "is": over facts, information, knowledge or beliefs about reality. These conflicts may center on the different perceptions of two people over basic assumptions. Opposition to a person's basic beliefs is a challenge to one's esteem. This fact explains, in part, why the mere presentation of the "true facts" of a situation to conflicting parties often does not resolve their differences. Evaluators must recognize that people give meaning to facts and that the "same" facts mean different things to different people.

(v) The Relationship Between the Conflicting Parties

People often conflict because of opposing views and desires in their relationship. Both may want to be "dominant" or "dominated" as the case may be. One may want more "closeness"; and so on. These conflicts over relationships often spill over into conflicts over issues which in some way involve the parties. Evaluators, in their attempts to help resolve "issues", therefore may discover that the real source of the issue lies in the biographies of the people associated with the problem and not the problem per se.

III. Charting The Variables Affecting The Course of Conflict

Conflict, its scope, level, intensity, duration, and mode of expression, will be affected by a number of complex variables. Accurate understanding of the context in which conflict exists, and particularly key variables which impinge upon it, is essential to evaluation. What evaluators must do, in a sense, is develop a category system for conducting an ethnography of conflict situations. As a step toward this objective, the following framework of variables provides a lens through which one can observe the dynamic forces likely to affect the course of conflict in an evaluative situation.

(1) Characteristics of the Parties in Conflict

a) Their Values and Motives

b) Aspirations and Objectives

c) Physical and Other Resources for Waging Conflict

d) Existing Beliefs About Conflict

* Ibid., pp. 16.

e) Perception of Reward/Gain - Punishment/Loss

f) Power Relationships: Equals/Unequals

g) Form: part vs. whole

(2) Prior Relationship of Parties to One Another

a) Attitudes and Beliefs About Each Other

b) Perception of Others View of Oneself

c) Degree of attitude Polarization: eg. good-bad, trustworthy-untrustworthy

(3) Characteristics of The Issue

a) Scope

b) Degree of Rigidity

c) Motivational Significance

d) Formulation

e) Periodicity

f) Consciousness of the Issue

(4) Social Environment of the Conflict

a0 Facilities and Restraints

b) Encouragements and Deterrents Re: Tactics and Strategies

c) Social Norms and Institutional Forms Re: Conflict Regulation

(5) Interested Audiences To The Conflict

a) Relationship to Parties in Conflict and to Each Other

b) General and Specific Interest in the Conflict

c) Potential Uses of a Resolved Conflict

(6) Consequences of Conflict To Participants and Others

a) Gains and Losses Relating to Issue

b) Precedents Established

c) Internal Changes in Participants From Engaging in Conflict

d) Long-term effects on Relationship of Parties Involved

e) Reputation of Parties in Eyes of Audiences

IV. Identifying Pre-Existing Processes and Procedures For Dealing With Conflict

All social contexts contain some mechanisms for dealing with conflict. Frequently, evaluators can use these pre-existing procedures for productive and co-operative conflict resolution. In some instances, it may only require the evaluator to bring these mechanisms to the attention of various parties involved. Two types of analysis, in particular, can be conducted: (1) Establishing Characteristics of the Strategies and Tactics available for Conflict Resolution (2) Assessing The Degree of Competition as Co-operation Within the Evaluative Context.

(1) Establishing The Characteristic of The Strategies and Tactics for Conflict Resolution

As in the case of describing the nature of conflict itself, conceptualizing the characteristics of available ways of resolving conflict requires an observational category system. A sample category system is presented below:

1. Relationship of strategies and tactics to:
   a) Assessing or changing one another's advantages and disadvantages.
   b) Influencing one anothers view of a) above
   c) Relative use of positive and negative incentives (eg. promises, rewards)
   d) Degree of free choice or coercion.
   e) Degree of information sharing.
2. Degree of credibility of existing processes and procedures
3. Degree of commitment of parties to resolution.
4. Types of motives appealed to.
5. Degree of stress on particular techniques, eg. coercion, persuasion, blackmail, seduction.

(2) Assessing The Degree of Competition and Co-operation

The core of the difference between co-operation and competition depends on the way in which the goals of the parties are linked. A co-operative situation is one in which the goals of the participants are so linked that any participant can attain his/her goal if, and only if, the others with whom he/she is linked cannot attain theirs. Each of the previous two descriptions, of course, present pure cases not often found in reality. In assessing co-operation and competition within an evaluative context, therefore, one is really looking for the degree of emphasis on a continuum.

ASPECT OF SOCIAL INTERACTION	COMPETITIVE PROCESS	CO-OPERATIVE PROCESS
I. Communication	1. Little communication or misleading communication. 2. Espionage or other techniques to obtain information the other is unwilling to give. 3. Each seeks accurate information about other, but to mistreat, discourage or intimidate other.	1. Open, honest communication of relevant information. 2. Main interest of each is to be accurately informed and to accurately inform. 3. Communications is persuasive rather than coercive in intent.
II. Perception	1. Increases sensitivity to opposed interests, to threats, and to power differences while minimizing the awareness of similarities.	1. Increases sensitivity to common interests while minimizing the salience of opposed interests. 2. A sense of convergence of beliefs and values.
III. Attitudes Toward One Another	1. Suspicious, hostile attitudes with readiness to exploit the other's needs and weaknesses. 2. Negative responsiveness to the other's requests.	1. Trusting friendly attitudes with a positive interest in the other's welfare. 2. Readiness to respond helpfully to other's needs and requests.
IV. Task Orientation	1. Stress on antogonistic interest. 2. Minimizations of the other's power of key objective.	1. Highlighting of mutual interests. 2. Co-ordinated effort with division of labor and specialization of function 3. Substitutability of effort rather than duplication.

By using specific aspects of social interaction, it is possible to identify focal points around which to develop criteria for determining the degree of competition or co-operation in a given context. The accompanying chart provides a shorthand method for such an assessment.

Conflicts, when placed in a co-operative context, become common problems in relation to which the conflicting parties possess a joint interest in reaching mutually satisfying solutions. A number of reasons can be offered as to why a co-operative process of conflict resolution in eavluation can lead to a constructive result:

(1) It aids open and honest communication of relevant information between participants.

(2) The freedom to share information enables the parties to go beneath the manifest issues to the underlying forces in the conflict and, thereby, facilitate a meaningful and accurate definition of the problems they are confronting together.

(3) It enables each party to benefit from the knowledge possessed by the other and thus to face the joint problem with greater intellectual resources.

(4) It reduces the likelihood of the development of a misunderstanding that can lead to confusion and mistrust.

(5) It encourages a recognition of the legitimacy of the other's interests and of the necessity to search for a solution that is responsive to the needs of each party.

(6) It tends to limit rather than expand the scope of conflicting interests and thus minimizes the need for defensiveness.

(7) It enables the participants to approach the mutually acknowledged problems in a way that utilizes their special talents and enables them to substitute for one another in their joint work so that duplication of effort is reduced.

(8) It leads to a trusting and friendly attitude which increases sensitivity to similarities and common interests.

As the foregoing indicates, co-operative conflict resolution through evaluation can exert positive impacts on communication processes, attitudes, perceptions and task orientations. Indeed, these alone are valuable outcomes of any evaluation.

V. Generating Alternative Strategies for Conflict Resolution

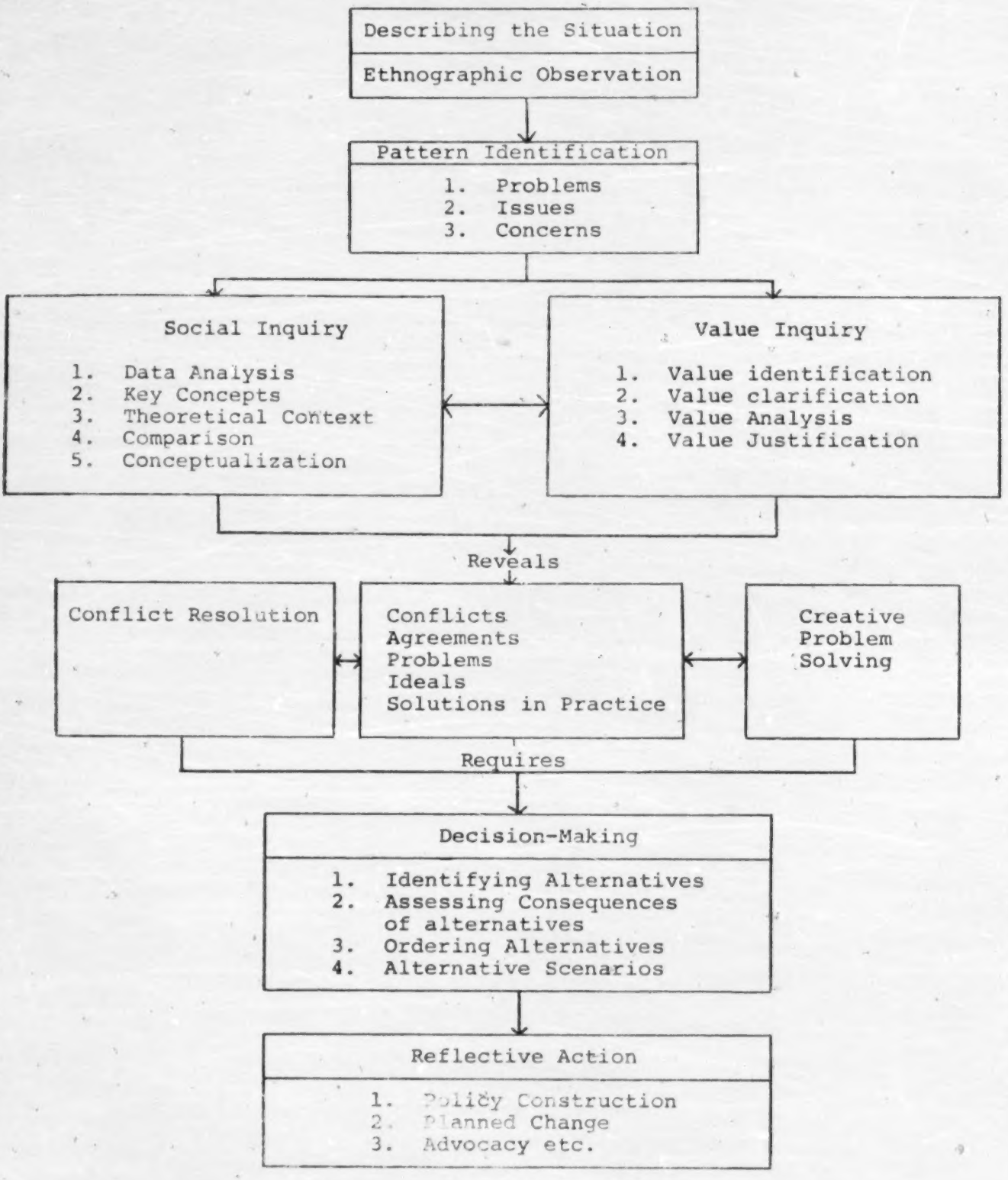
Evaluators, in attempting to deal constructively with conflict, must be wary of falling into the trap of forcing that conflict into a pre-conceived mould for resolution. Flexibility and a willingness to "roll with the punches" best characterize the evaluator's stance in relation to conflict. More often than not, in fact, evaluators will increase their effectiveness (and degree of adoption of their proposals) if they acknowledge and experiment with forms of conflict resolution already existing in the situation. The choice as to which form will best facilitate resolution—be that through the use of such institutional forms as collective bargaining, social roles like referees or concilators, social norms like fairness or rules governing the interaction and negotiation—will depend largely upon the evaluator's sensitivity to and understanding of the dynamics inherent in the situation. It is at this point, in other words,

that the prior recognition of value conflict and the use of ethnographic techniques become invaluable.

Conclusion

This paper has attempted to offer a critique of conventional approaches to evaluation in educational settings as well as to sketch out the form of an alternative strategy; one that is value-based in nature, developmental in purpose and experiential in its process. Further refinement of the concept and a serious attempt at experimentation and training will be required if the approach outlined is to exert a general impact on education. The diagram below represents an attempt to capture and link the components of the approach within a working framework for evaluation. The framework rests upon several postulates concerning the nature of evaluation. These are presented as an agenda for those concerned about re-thinking the concept, role and process of evaluation in education.

An Experiential Model For Evaluation



An Agenda For Evaluation Reform

I The conduct of Evaluation is a cultural act rooted in social interaction.

II Each participant in an evaluation process responds to others in terms of his/her perceptions and ideas of the other; these may or may not correspond to the other's reality.

III Each participant in an evaluation process, being aware of the other's capacity for awareness, is influenced by his/her own expectations concerning the other's actions as well as by his/her perception of the other's conduct.

IV Evaluation, as a process of social interaction, is not only initiated by values and motives, but generates new motives and values and alters the old ones. "It is not only determined but also determining."

V In the process of rationalizing and justifying actions that have been taken and effects that have been produced, evaluation allows new values motives to emerge.

VI Decision-making in evaluation can entail a struggle among different interests, values and ideals for control over the process.

VII Evaluation takes place in a social environment—in a family, a group, a community, an institution, a civilization—that has developed techniques, symbols, categories, rules, and values that are relevant to the effectiveness of the evaluation. To understand the events that occur as part of an evaluation process, one must comprehend the interplay of these events with the broader social context in which they occur.

VIII Evaluation is a prescriptive activity in which the evaluator bears part of the responsibility for the initiation of appropriate action.*

*Details of this approach can be found in T. R. Morrison and J. A. Riffel, Evaluation and Co-operative Conflict Resolution in Education (unpublished manuscript, 1975)

TOM MAGUIRE

The problem, it seems to me, is that both Tony and Terry are pushing a point of view that (a) values are central and (b) it is common that you get into values conflicts very quickly. I agree that values are central to evaluation, but I disagree with the notion that you very quickly get into disagreements, into values conflicts. What you get into are conflicts with respect to goals. On the basis of this morning's experience it seems to me that where we disagreed was on what the goals were. We would come, I think, at some level of generality - perhaps a reasonably high level of generality - to agree that our values were approximately the same. We'd rather have people alive than dead, we'd rather have people happy than sad. We would agree, and there would be very little values conflict. The problem comes in what should be the appropriate goals in a situation and I think that's quite a different kind of argument. If the values are reasonably common, then we can say that for these goals these kinds of procedures will fit. What I'm beating around the bush saying is I'm not terribly optimistic about the resolution of values conflicts.

Question: I think I would like to ask Tom about the goals aspect of it. We're talking about an educational context and you're suggesting there will be goals conflicts. What reaction do you have to models like the one that's disseminated by Phi Delta Kappa, in which the people decide the goals?

McGuire: Let me try to answer it by what I think the evaluator should do. One thing an evaluator should do is propose goals that are alternatives to the goals that are being espoused by whomever. The fact may be that they are equally valid in some sense and that indeed the kinds of processes that are being pursued are more adapted to these other sets of goals.

Let me go back. It seems in most educational situations we sort of go on and on and on. We never really decide, despite the harping, what it is we're trying to do. We just continue on doing it and we hope that we're iterating toward something. Some people would say that we've been beating around this goals definition thing for a long time. It seems to me that we haven't been very successful about it. We should constantly, I think, be harping on an explication of goals because it may turn out that the goals that we say we're pursuing and those that we are observed by others to be pursuing are quite disparate. And I'm reasonably confident that I have avoided answering your question and that's what the problem is, because your goal and my goal were quite separate in that.

Question: I'd like to put a question to Tom. I always think of an evaluator as being one who is checking to see whether or not goals are being attained. Now you mention that an evaluator should also propose alternative goals. Now, are these two not in conflict or are you broadening the definition of how I have seen an evaluator?

McGuire: No, I don't know that they're in conflict. What I've observed is that people take goals for granted. Goals become self-evident to them. And off they go and pursue them. One thing that an evaluator can do is have the clientele look and say "Yes, these really are the goals" and balance whatever the outcomes are with the goals. One must then be prepared to say "Well, those really weren't the goals to begin with" and change one's intentions to match the outcomes. But very often that hasn't been done because of the way we oepr-ate in school systems. We just kind of flow from year to year. We think we're operating in a well-defined situation. The goal may be to make these kids learn better, but nobody ever thought about the notion of the goal being to make kids happier, I think that one thing evaluators can do (and the more we talk about it the less I like the notion of evaluators) is to help people look at kinds of goals that are appropriate in a situation. I don't feel a great deal of conflict there.

Question: I'm bothered by what you say. If the expectation of the evaluator, whatever he's called, is that he will help in goal clarification, I think that is presumptuous, an intrusion on his part regarding what those people need. He doesn't know them particularly; he doesn't know their program; but the expectation that they're going to need some goal clarification is a consider-able presumption.

McGuire: That's right. And that's why I'm bothered with the notion of the evaluator as being that person who comes in from the outside. I'm much more comfortable with a model that I thought was evolving from Tony, where I say a function being set up like a decision-facilitator, who would have certain kinds of special skills. I don't think goal-clarification would be presumptuous for that kind of person.

There is another function and it's the function that I'm more interested in. It has to do with collecting data that would be useful for this group of people that has the mandate, however they got the mandate. This presumptuousness, it seems to me, is a valid charge when we're talking about the sort of thing that I used to push for - and that is the all-knowing, Mr. Clean evaluator. You see, the notion I used to have was that what we would do in a particular province, would be to bring all the evaluators into a central checking place and test them to make sure that they were clean-in an evaluative sense. We would stamp them as evaluators and they would be allowed to go out and practice evaluation for a year. Then it occurred to me that you can get contaminated very quickly, so you might have to have this pro-cess going on weekly and then it jsut all falls apart for me. We just can't go with that model. I'm glad that you're uncomfortable with that.

It seems to me that formal models do have utility. Unfortunately they don't have the utility that I used to think they had. The utility seems now to me to have something to do with providing suggestions and directions for data collection. Last night I heard somebody at the back of the room say that the techniques of evaluation were pretty well advanced and that where we should spend our time is in looking at the problems of decision-making. But I think that's where the problems are. The problems are on the other side with the techniques of evaluation, the collection of evidence that has some kind of reflection on the truth of the situation. We should be spending

more time developing our notions about data collection. This is where I see the utility of formal somethings, models, procedures, as coming into it.

When you ask the question whoever used these models and you search around, you find out that here and there there is an example of a model being applied. But most often when you talk to people who do evaluations you find them pretty eclectic...a little bit from here, a little bit from there. They used the models to stimulate their minds to decide what information to collect. There never were many applied models.

Question: I just wondered if we can go back to the first question? I don't know what we're talking about when we're talking about the evaluator. I suspect that we're talking about quite separate things. I had a picture a few years ago about a person who comes in and evaluates something and that really was one person, commanded the data collection, did the suggesting, produced the recommendations, the whole business. If we're talking about how deeply that person **should** be involved then the question doesn't make sense because I don't think that person exists any more. Probably never did.

McGuire: Evaluation goes on in a kind of committee system. The things that count where the decisions are going to be made, that goes on in committee. Now, if we're talking about people on those committees, in a group that is trying to reach a decision; if we're calling those the evaluators and you say to me, how will we resolve conflicts between what people consider worth-while within that committee, then I say to you in my role as a person who used to be labelled evaluator, I have no special knowledge to deal with that. I used to think I did but I don't think so any more. What suddenly dawned on me was that I had been thinking that when people were talking about value conflicts, we were talking about behavioral dispositions, as opposed to assessments of things that are the worth of objects or situations. I'm not really, in these committee settings, optimistic about resolving different value stances that people might have; these have to do with their personality and their approach to situations. I am a little more optimistic about debating for example, in a particular setting,"Is it better to have kids happier or better spellers?" My problem is that I'm a poor Prairie-cum-postal boy who is at the concrete operational stage and when I deal with evaluations they're simple kinds of problems. Say we've got a program and everybody's happy about it, except the kids can't spell very well. And the evidence supports that contention. Simplistically the decision seems to me, should we have happy poor spellers or unhappy better spellers. What I'd hope is that any sort of any sensible group would look at the range of possibilities and explore the consequences of various lines of action.

Question: Good. Now that's what I see the role of this person, whoever it is, as being.

McGuire: And I do not. I don't think the evaluator knows anything about that sort of raising of options. The evaluator as a person, as a specialist, usually goes into the situation very unacquainted with what the options are, and spends a lot of time trying to learn them from the people.

Question: May I ask a question? Is this sort of what you have in mind...an evaluation team coming into, say, a school situation and making an evaluation of a particular program?

McGuire: No. Not any more. It's a group of people, maybe from within the school, who says to themselves, "We would like to know what kind of job we're doing." That's the kind of group I am talking about. Where they need help is at two points. They need help from a facilitator of some sort to help chair the thing to help draw out whatever, and they will need expert help in data collection.

* * * * *

BOB STAKE

Last night Tony gave a central place to the responsibility for value clarification in the evaluation process. I do not. I agreed with him on many of the points that he made in his paper last night but not on that one. If I was favored with our usual instrument, the blackboard, I would put up a mnemonic which would end with the question of whether or not the action phase belongs within the evaluation responsibility, including the making of recommendations. Though I myself many times make recommendations and many times I am obligated by contract to make recommendations, I feel that I am very likely overstepping the bounds of the evaluation responsibility and moving into the decision-making responsibility. Value clarification or value resolution or conflict resolution is a matter I also consider to be part of the action phase and outside the evaluation responsibility. I do not see consensus or the resolution of such differences as may be found as being within the evaluation responsibility.

Now that, I think, needs some clarification. As Tom has just said, it is clear that evaluation is concerned with values. Whether it's going to deal with them in an explicit form or not is a matter of strategy, a matter of negotiation, a matter of considerable importance. It should not be concluded automatically that because values are a central concern to evaluation they should be explicated, studied, dealt with in some sort of resolution form. Maybe they should and maybe they shouldn't. One cannot carry out the data gathering, presentation and analysis, phases of the evaluation without dealing with values in implicit ways. I'm much in favor of dealing only with the intrinsic values and not putting ourselves into a constructed evaluation setting as we were this morning. The evaluator has an obligation to find out what intrinsic valuing there is in the situation; for example, how different groups of teachers are valuing things, how individuals - students, administrators, other of importance - are valuing the program already, or valuing new reflections that are brought by the evaluator. To set up a situation where there is an examining of the values explicitly involves a constructed as opposed to intrinsic set of values, and I'm not at all enthusiastic about it.

There will be value conflicts, of course, but the resolution of value conflicts is not an essential part of the evaluation responsibility.

Question: Could I ask Bob Stake to say a bit more about the distinction I understood you to have made between evaluation and action?

Stake: For the purposes of this mnemonic I would say that there is a structuring phase, there is a data-gathering or observation phase, there is a presentation or incoding phase, and then there is a validation or triangulation phase or some way of substantiating what the presentation is. That constitutes the responsibility of the formal evaluation study. Perhaps in the informal

evaluations that we're doing regularly one could see some of that same sort of implicit activity but there it stops. The idea that evaluation should be an integral part of decision making, that if there is a person sho is an evaluator that person should share in the decision-making responsibilities is drawing the boundary for evaluation as a function beyond what it should be. I'm not saying that in some situations the evaluator should't help out and be the decision-maker. Evaluation in many educational situations is seen as the responsibility for providing a situation in which the people normally responsible for taking action and understanding the program can do their research and reflection better. It's like a teacher in a discovery learning situation. The evaluator provides a research scene for people who ordinarily are responsible for a program to understand that program better. Nothing very much channels through the evaluator's head; the quality of the evaluation is carried on quite outside the evaluator's head. He's an arranger of things. Sure, he stimulates a little bit, but the decision - the action process - is really quite outside his territory.

I have made quite a big fuss the last two or three years. Some of you might have read some things I have written in which I make a big contrast between the preordinative approach to evaluation and the responsive approach to the evaluation, with a strong personal committment to the latter. That might appear to mean that I don't believe in the a priori model, but what I am opposed to really is the presumption that the key issues in an educational situation are apparent either through statements of objectives or through some kind of posing of a problem. The key things to be gathering data on and the issues that are of most concern to people often will lie pretty well hidden, and one needs to have considerable contact with those programs before being able to recognize those issues. The a priori part of the responsive approach is one of putting oneself in an observation situation in contact with a great number of different audiences and a number of scheduled interactions, of anticipating certain things.

Question: I just wondered in connection with the first question if we could go back to that momentarily. It seems to me that the panel didn't really come to grips with the first part of that question, which was if values are central to the evaluation.

Stake: They're central. You may choose to deal with them in an explicit sort of way, where you set up a constructed situation. Or you may just leave them in the intrinsic situation and not talk about the extraction of values, but what sorts of action, what sorts of negotiation or what might be done about it.

Question: Are you saying then that it's not the role of an evaluator to be involved in resolving value conflicts which do surface during the act of evaluation?

Stake: I'd say no. That might be the right thing for that person who beside being an evaluator is an educator, is a human being, is a friend; it might be the right thing for that person to do but it doesn't fall within the brief to do it.

It is important to find out what sorts of conceptualizations are in each of our minds as we're talking. But I do not really believe that it makes that much difference. I believe this whether you are talking about a single person exercising many many diverse responsibilities for education, one of which is evaluation, or you are talking about a very large team assembled to evaluate the entire public education system in this Province. Once you get the territory in mind then you start going back to what are the degrees of using a priori model or how deeply to become involved. The principles, the rules, I think, are pretty general. Certain things happen if you presume that a pre-test, post-test model is going to work for you, or if you presume that knowledge of the kind that a scientist collects when he uses an interview schedule of a certain kind does one thing for you but not something else. It seems to me that the answers come out about the same regardless of the concretization that we assume.

I have a feeling that we are talking about a very limited range of evaluation. If you think of the whole spectrum our conversation seems to be particularly related to the evaluation of programs. There is a whole range of other things that we might be evaluating - economic costs, for example.

The example given about knowing the effect of certain textbooks and the like; I don't see that as evaluation when you specify the criterion which you did. But if you ask, "What are the effects of these textbooks?", then that is an evaluation question. If you ask what are the effects of the community response now to the latest request for more money by the school board, these are evaluative questions that call for weighing the desirability of various positions and the like, so I include those within the sense of program evaluation, though I'm usually talking about something that has a better definition that that...I'm certainly not talking about myself. I'm not talking about an act, like the act of the skipper this morning in deciding that certain people have to be cut away. I'm <u>not</u> talking about that kind of evaluation; that <u>is</u> another important kind.

Nine Approaches to Educational Evaluation — prepared by R. Stake, CIRCE, September, 1974

APPROACH	PURPOSE	KEY ELEMENTS	PURVIEW EMPHASIZED	PROTAGONISTS (see references)	CASES, EXAMPLES	RISKS	PAYOFFS
STUDENT GAIN BY TESTING	*to measure student performance and progress*	goal statements; test score analysis; discrepancy between goal and actuality	EDUCATIONAL PSYCHOL-OGISTS	Ralph Tyler Ben Bloom Jim Popham Mal Provus	STEELE WOMER LINDVALL-COX HUSEN	*oversimplify educ'l aims; ignore processes*	*emphasize, ascertain student progress*
INSTITUTIONAL SELF-STUDY BY STAFF	*to review and increase staff effectiveness*	committee work; standards set by staff; discussion; professionalism	PROFESSORS, TEACHERS	National Study of School Evaluation Dressel	BOERSMA-PLAWECKI KNOLL-BROWN CARPENTER	*alienate some staff; ignore values of outsiders*	*increase staff awareness, sense of responsibility*
BLUE-RIBBON PANEL	*to resolve crises and preserve the institution*	prestigious panel; the visit; review of existing data & documents	LEADING CITIZENS	James Conant Clark Kerr David Henry	FLEXNER HAVIGHURST HOUSE ET AL PLOWDEN	*postpone action; over-rely on intuition*	*gather best insights, judgment*
TRANSACTION-OBSERVATION	*to provide understanding of activities and values*	educational issues; classroom observation; case studies; pluralism	CLIENT, AUDIENCE	Lou Smith Parlett-Hamilton Bob Rippey Bob Stake	MacDONALD SMITH-POHLAND PARLETT LUNDGREN	*over-rely on subjective perceptions; ignore causes*	*produce broad picture of program; see conflict in values*
MANAGEMENT ANALYSIS	*to increase rationality in day to day decisions*	lists of options; estimates; feedback loops; costs; efficiency	MANAGERS, ECONOMISTS	Leon Lessinger Dan Stufflebeam Marv Alkin Alan Thomas	KRAFT DOUGHTY-STAKENAS HEMPHILL	*over-value efficiency; undervalue implicits*	*feedback for decision making*
INSTRUCTIONAL RESEARCH	*to generate explanations and tactics of instruction*	controlled conditions, multivariate analysis; bases for generalization	EXPERIMEN-TALISTS	Lee Cronbach Julian Stanley Don Campbell	ANDERSON, R. PELLA ZDEP-JOYCE TABA	*artificial conditions; ignore the humanistic*	*new principles of teaching and materials development*
SOCIAL POLICY ANALYSIS	*to aid development of institutional policies*	measures of social conditions and administrative implementation	SOCIOL-OGISTS	James Coleman David Cohen Carol Weiss Mosteller-Moynihan	COLEMAN JENCKS LEVITAN TRANKELL	*neglect of educational issues, details*	*social choices, constraints clarified*
GOAL-FREE EVALUATION	*to assess effects of program*	ignore proponent claims, follow check-list	CONSUMERS; ACCOUNTANTS	Michael Scriven	HOUSE-HOGBEN	*over-value documents & record keeping*	*data on effect with little co-option*
ADVERSARY EVALUATION	*to resolve a two-option choice*	opposing advocates, cross-examination, the jury	EXPERT; JURISTIC	Tom Owens Murray Levine Bob Wolf	OWENS STAKE-GJERDE REINHARD	*personalistic, superficial, time-bound*	*info impact good; claims put to test*

Of course these descriptive tags are a great over-simplification. The approaches overlap. Different proponents and different users have different styles. Each protagonist recognizes one approach is not ideal for all purposes. Any one study may include several approaches. The grid is an over-simplification. It is intended to show some typical, gross differences between contemporary evaluation activities.

ALKIN, Marvin C. Towards an evaluation model: A systems approach. UCLA: Center for the Study of Evaluation, Working Paper #4, December, 1967.

ANDERSON, Richard C. The comparative field experiment: An illustration from high school biology. Proceedings of the 1968 Invitational Conference on Testing Problems. Princeton, New Jersey: Educational Testing Service, 1969, 3-30.

BLOOM, Benjamin S., Hastings, J. Thomas, and Madaus, George F. Handbook on formative and summative evaluation of student learning. New York: McGraw-Hill, 1971.

BOERSMA, Wendell C. and Plawecki, Henry M. After the NCA self-evaluation and visitation--what happens? North Central Association Quarterly (5454 South Shore Drive, Chicago), Winter, 1972, XLVI(3), 335-339.

CAMPBELL, Donald. Reforms as experiments. American Psychologist, 1969, 24, 409-429.

CARPENTER, James L. Accreditation evaluation and institutional change. North Central Association Quarterly (5454 South Shore Drive, Chicago), Fall, 1970, XLV(2), 259-263.

COHEN, David K. Politics and research--evaluation of large-scale programs. Review of Educational Research, 1970, 40(2), 213-238.

COLEMAN, James S. Policy research in the social sciences. Morristown, New Jersey: General Learning Press, 250 James Street, 1972.

COLEMAN, James S. et al. Equality of educational opportunity. Washington, D.C.: U.S. Government Printing Office, Supt. of Documents, 1966.

CRONBACH, Lee J. Course improvement through evaluation. Teachers College Record, May, 1963, 64, 672-683. Reprinted in R. W. Heath (Ed.), New Curricula, 1964.

DOUGHTY, Philip L. and Stakenas, Robert G. An analysis of costs and effectiveness of an individualized course in college level geology. In Creta Sabine (Ed.), Accountability: Systems Planning in Education. Homewood, Illinois: ETC Publications, 1973.

DRESSEL, Paul L. Accreditation and institutional self-study. North Central Association Quarterly (5454 South Shore Drive, Chicago), Fall, 1971, XLVI(2), 277-287.

FLEXNER, Abraham. Medical education in the United States and Canada. A report to the Carnegie Foundation for the Advancement of Teaching. New York: The Carnegie Foundation, 1910. Reprinted, New York: Arno Press, 1972.

HAVIGHURST, Robert J. The public schools of Chicago. A survey for the Board of Education of the City of Chicago, 1964.

HEMPHILL, John K. The relationships between research and evaluation studies. In Ralph W. Tyler (Ed.), Educational Evaluation: New Roles, New Means. Chicago: University of Chicago Press, 1969, 68th Yearbook, Part II, National Society for the Study of Education, 189-220.

HOUSE, Ernest, Rivers, Wendell, and Stufflebeam, Daniel. An assessment of the Michigan accountability system. The authors: March, 1974.

HOUSE, Ernest R. and Hogben, Donald. Setting speculative shares. SRIS Quarterly, Spring, 1973, 6(1), 11-13.

HUSEN, Torsten. International study of achievement in mathematics. New York: Wiley, 1967.

JENCKS, Christopher et al. Inequality. New York: Basic Books, 1972.

KNOLL, Robert E. and Brown, Robert D. Experiment at Nebraska: The first two years at a cluster college. University of Nebraska Studies, new series, #44. Lincoln: University of Nebraska, June, 1972.

KRAFT, Richard H. P. Cost-effectiveness analysis in vocational technical education. AERA Monograph Series on Curriculum Evaluation, Volume 7. Chicago: Rand McNally, 1974.

LESSINGER, Leon M. Accountability for results: A basic challenge for America's schools. In Leon M. Lessinger and Ralph W. Tyler (Eds.), Accountability in Education. Worthington, Ohio: Charles A. Jones, 1971.

LEVINE, Murray. Scientific method and the adversary model: Some preliminary suggestions. Evaluation Comment, 1973, 4(2), 1-3.

LEVITAN, Sar A. Antipoverty work and training efforts: Goals and reality. Policy papers in Human Resources and Industrial Relations, Number 3. Ann Arbor: Institute of Labor and Industrial Relations, August, 1967.

LINDVALL, C. M. and Cox, R. C. Evaluation as a tool in curriculum development: The IPI evaluation program. AERA Monograph Series on Curriculum Evaluation, Volume 5. Chicago: Rand McNally, 1970.

LUNDGREN, Ulf P. Frame factors and the teaching process. Stockholm: Almquist and Wiksell, 1972.

MACDONALD, Barry. Humanities curriculum project. In Schools Council Research Studies, Evaluation in Curriculum Development: Twelve Case Studies. London: Macmillan, 1973, 80-90.

MOSTELLER, Frederick and Moynihan, David P. (Eds.) On equality of educational opportunity. New York: Vintage Books, 1972.

NATIONAL Study of Secondary School Evaluation. Evaluative criteria, fourth edition. Washington, D.C.: National Study, 1785 Massachusetts Avenue, N.W., 20036, 1969.

OWENS, Thomas R. Educational evaluation by adversary proceedings. In Ernest House (Ed.), School Evaluation--The Politics and Process. Berkeley: McCutchan Publishing Corporation, 1973.

PARLETT, Malcolm. Study of two experimental programs at MIT. In David Hamilton, Malcolm Parlett, David Jenkins, and Barry MacDonald (Eds.), _____ A Reader in Unconventional Curriculum Evaluation. New Education Press, in press.

PARLETT, Malcolm and Hamilton, David. Evaluation as illumination: A new approach to the study of innovatory programs. Centre for Research in the Educational Sciences, Occasional Paper 9. Edinburgh, Scotland: University of Edinburgh, October, 1972.

PELLA, M., Stanley, Julian C., Wedemeyer, C. A., and Wittich, W. A. The uses of the White films in the teaching of physics. Science Education, 46, 6-21.

PLOWDEN, Lady Beatrice et al. A report of the central advisory committee on children and their primary schools. London: Her Majesty's Stationery Office, 1967.

POPHAM, W. James. Objectives and instruction. In Robert Stake (Ed.), AERA Monograph Series on Curriculum Evaluation, Volume 3, Instructional Objectives. Chicago: Rand McNally, 1969, 32-64.

PROVUS, Malcolm. Discrepancy evaluation. Berkeley: McCutchan, 1973.

REINHARD, Diane L. Methodology development for input evaluation using advocate and design teams. PhD dissertation, Ohio State University, 1972.

RIPPEY, Robert (Ed.) Studies in transactional evaluation. New York: McCutcheon, 1973.

SCRIVEN, Michael. Prose and cons about goal-free evaluation. Evaluation Comment, December, 1972. Also in W. James Popham (Ed.), Evaluation in Education: Current Applications. Berkeley: McCutchan, 1974.

SMITH, Louis M. and Geoffrey, L. The complexities of an urban classroom: An analysis toward a general theory of teaching. New York: Holt, Rinehart, and Winston, Inc., 1968.

SMITH, Louis M. and Pohland, Paul A. Education, technology, and the rural highlands. AERA Monograph Series on Curriculum Evaluation, Volume 7. Chicago: Rand McNally, 1974.

STAKE, Robert E. and Gjerde, Craig. An evaluation of TCITY, the Twin City Institute for Talented Youth. AERA Monograph Series on Curriculum Evaluation, Volume 7. Chicago: Rand McNally, 1974.

STAKE, Robert E. The measuring of education. Berkeley: McCutchan, in press.

STANLEY, Julian C. Controlled field experiments as a model for evaluation. In Peter H. Rossi and Walter Williams (Eds.), Evaluating Social Programs. New York: Seminar Press, 1972, 67-71.

STEELE, Joe M. Arranging field tests: Characteristics of sites and students. Formative Evaluation Report No. 1. Boulder, Colorado: Biological Sciences Curriculum Study, June, 1973.

STUFFLEBEAM, Daniel L. et al. Educational evaluation and decision-making. Itasca, Illinois: Peacock, 1971.

TABA, Hilda. Teaching strategies and cognitive functioning in elementary school children. Cooperative Research Project No. 2404. San Francisco: San Francisco State College, February, 1966.

THOMAS, J. Alan. Cost-benefit analysis and the evaluation of educational systems. Proceedings of the 1968 Invitational Conference on Testing Problems. Princeton, New Jersey: Educational Testing Service, 1968.

TRANKELL, Arne. Kvartet Flisan. Stockholm, 1973.

TYLER, Ralph W. Basic principles of curriculum and instruction. Chicago: University of Chicago Press, 1950.

WEISS, Carol H. Evaluation research: Methods of assessing program effectiveness. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1972.

WOLF, Robert L. The application of select legal concepts to educational evaluation. Unpublished PhD dissertation, University of Illinois, 1974.

WOMER, Frank B. et al. National assessment of educational progress, report 1, 1969-70 science: National results. Washington, D.C.: U.S. Superintendent of Documents, July, 1970.

ZDEP, Stanley M. and Joyce, Diane. The Newark-Verona plan for sharing educational opportunity. PR-69-13. Princeton, New Jersey: Educational Testing Service, September 1, 1969.

LIONEL ORLIKOW

I'm not quite sure what expectations have been built up but as this is a session on evaluation I'm sure that whatever your models are, you'll be acting and re-acting quite quickly in the next half-hour. The topic that I have is Dilemmas in the Evaluation of a Provincial Educational System. I believe that my task is to provide some of the background of where the Department has been and where it will go, and give you an opportunity to look at a number of the variables that we've found in trying to think through an evaluation system for the Province. I apologize now - this will be my last apology - and my first - because there will be a number of generalizations made. If I had time, and perhaps inclination, I would try to cover my tracks, so that the exceptions, the modifications, would be recognized. But I will make generalizations deliberately in order to enforce a particular point of view.

The point in my introduction that I would like to stress is that evaluation is political as all educational issues are political. Unfortunately, most educators refuse to look at the political environment which permeates all aspects of education. As that is rate in education - thinking of power, resources, conflict, compromise - it makes it very difficult to look at questions such as evaluation, in a proper spirit. I would also like to stress in my quick general introduction that I am a person who has to make decisions. So when I look at the dilemmas which surround evaluation I am in an unfortunate position of having to make decisions about reconciling a number of the very discordant cries that come in from the field to keep the operation going. As well, education is an operation where there are a host of variables that all have to be juggled at the same time and I think it's quite obvious that in juggling a large number of eggs, some of them fall and break. The dilemma issues which I have cited: central, local, the management, the diagnostic, the research, the practice - continue to bedevil me. However, I think that you'll see the themes come through in the next half hour.

I'll start off by saying that as far as we in the Department are concerned there is a tremendous amount of information available on evaluation but we cannot sit about and debate how many evaluators could dance on the head of a pin. We have to proceed to look at a number of steps in evaluation. What my talk will attempt to do is first to describe the major activities in evaluation within the department; secondly, to outline in clear detail why this particular direction has been undertaken, and then to look at the overall climate as I see it in education - the teaching situation and its relevance to evaluation, and a number of Manitoba variables that assist and retard in the application of any evaluation system. I will conclude by sketching out what I think is the basic issue related to evaluation, why it is ignored and why I am trying to keep evaluation at a moderate level.

Now the first point, what has the Department done? The Department of Education in Manitoba in the last few years has moved very slowly in the evaluation field; much more slowly than a large number of individuals would desire. There are a number of teachers and University professors for example who wish to return to the good old days of the high school departmental examinations. There are a rising number of school trustees who wish that standardized tests be applied in order to evaluate and assess school divisions, schools, and teachers. Over the last 3 years a variety of alternatives have been examined and on the whole discarded. We have looked at a variety

of State Goals Exercises that have been done in the United States and we have rejected them. While they sound very good on the surface, the impossibility of getting large numbers of the public involved in what I consider to be theoretical exercises, the translation of these theoretical goals into actual application within schools I think is a will-o'-the-wisp. We have looked at the National Assessment of Educational Progress in Denver and while there are certain instruments there that are of some interest to us we feel that the data is too gross and has little diagnostic use in schools or school divisions. Criterion Referenced Tests are very interesting. However, we have very limited resources and I'm not too sure if we pumped in the large amount of resources needed for criterion referenced tests whether there would be very much transfer into the teaching situation. PPBS, MBO: very interesting again but let us leave them with the United States Defence Department, and not apply them in education. So in looking at a numbered of the more established, the more popular types of evaluation systems throughout the North American continent what have we developed?

Currently, our greatest hope is in the External-Internal model which is being tested in three school divisions of the province. The External-Internal model is somewhat similar to the OECD reviews that are introduced where an external team visits a particular locality and the internal group also have made an assessment, a self-examination of their own performance. There is a similarity in what is happening in Ontario at the present time where they are using some sort of external-internal mode, except we believe the Ontario Department of Education is stressing the external whereas we are attempting to stress the internal. The steps within the model are probably quite familiar: the derivation of the goals, the identification of the needs and then the consequent program responses. How there is an urgent need within this province for a variety of reasons to have some form of a provincial review on all school divisions done on a regular basis, every five years or six years. The external-internal model does provide a balance between the needs of the local school division and the provincial imperative in education. There are a number of obvious problems in resource allocation, on what are the provincial goals and so on.

Again I'd like to stress that one of our main problems in developing an evaluation model in this province is not to get the ideal model but one which I would call in economic terms 'satisficing'. In view of the many contraints and restraints what is the minimum that will do the greatest amount within a very, very constrained environment? At the school level there have been two operations that have been conducted. One is the community assessment program which was conducted in a few schools in the inner core of Winnipeg. Basically, when the activities are looked at over the last year, it turned out to be more of a community participation model, looking at all activities within a school rather than evaluation in even the very general sense of the term. I believe that the community assessment program is a successful model but unfortunately it is also a failure. And I think that the failure overwhelms the success. That is, to introduce the model is very expensive and very time-consuming. What can happen within three to four schools is impossible for a provincial system; unless the Province has the money - which it does not - to put in a considerable amount of funds.

Another evaluation device that has been used is questionnaires with individual business concerns and parents of children who are in experimental or open schools. The conclusions of the evaluations of educational programs through external people have been quite interesting in that they have completely contradicted the findings that are repeated consistently through the mass media, namely that there is a high degree of dissatisfaction within business and from parents about the education of their children. For example, the Canadian Chamber of Congress has had considerable play in the press over their concern that children coming into their employment have not got the skills and discipline - they can't read, they can't write, and so on. Now aside from the fact that the instruments used by the Canadian Chamber were very bad, the main finding that we had was that while people are prepared to decry education on the global level, when the individual businesses were asked about the success of children who were in their own particular factory or office the response was overwhelmingly positive. Unfortunately, while I think our results are very interesting, I'm afraid that the informal evaluation as portrayed through the mass media has completely overwhelmed us and the truth has not come out.

A third area that the province is going to become involved in, and I do this very reluctantly, is a Provincial form of standardized testing. I'm trying to avoid it for many reasons-which will become apparent later but I do believe that there is need for some form of gross data to provide information on trends over time as well as a form of data to deal with concerns that children are not reading or writing properly. Also the trustees of the province do need some information. In their particular situation, where they have to get elected, one of the frequent charges made against them is that the children are going to rack and ruin. In the development of our testing strategy there will have to be considerable compromises made because I do not, and I stress that, I do not and I will not countenance any form of testing scheme that will be applied to compare teachers and compare schools. I hope that we will also be able to avoid comparisons among school divisions.

What have been the criteria that we have been looking at in the development of the scheme? First, I think it has become quite apparent that whatever scheme is developed must be simple and easily understood, by teachers as well as parents. The second criterion is that whatever system is developed must be relatively easy to establish, and I mean ease in terms of staff time, dollars, and other types of resources. The third is that the modicum of flexibility that has been introduced within the provincial education system over the last decade has to be preserved - hopefully extended, but I think in this day, preserved, and not constrained. And fourth and here we have had I would say on the whole very little success - there has to be complete participation as far as possible by teachers and parents. I don't mention students, because we've had no success there at all.

So I start off with the basic problem that we've had. We're starting with what seems to be a very, very simple truth. There is a need for basic data. Everyone wants to improve decision-making. And that goes at all levels of education, from the classroom even to 1181 Portage. That might sound a contradiction, I know, but I throw that out. In all the shenanigans in developing these very elaborate strategies, I think we would all agree, although we sometimes forget, that this will lead to better educational environments for students. Somehow I think they get lost in many of our adult concepts of reality.

In trying to look at the dilemmas that we've had, I'd like to look at a few of the items in the environment in which I have to move at the present time. The first is the over-all mood within education today. I believe it is a well-known aspect of education that we as educators are very much influenced by fads, and I do believe that evaluation has to be placed in terms of some of the faddism in which we are existing today. I would say that the pendulum is swinging to a very centralist and a very narrow definition of education. We see the usual questions. There must be a return to the basics. (I didn't know we ever left them.) We see a number of our friendly educators, that were the pseudo-progressives of the late sixties, the early seventies, now becoming the crypto-conservatives of the late seventies. The rhetoric has changed in the last five years among a number of our education leaders. In fact, for those of you who taught in the late nineteen-forties I think you'd find many of the themes very familiar. However, we do not deal with these types of very strong value issues. We try to duck them.

At present we also have the growth of our management people who wish to put everything into good clear terms. Let us reduce waste. Let us make sure that those progressives who've gone wild have some control. The objectives managers are becoming more popular, in trying to bring some type of order into our system. I believe that the swing of the pendulum, as supported by a number of the enthusiasts in evaluation, is going to strengthen two of the main aspects of our educational system that have prevented very much flexibility. I could see a fair amount of the evaluation done reinforcing the centralists. And I say the centralists not just in terms of those who see the Department of Education being able to levy a uniform system across the province of Manitoba, or those at the school division level who also have little trust, little confidence in their classroom teachers and wish to control them from some central office. I also see the technocrats in education, the narrow specialists who rely now not just on the maths experts to come in but also the evaluation expert to come from the central office to provide a form of expertise that will further limit the discretion of the teacher in adapting his or her classroom environment to the needs of children. Evaluation is perfect in meeting the needs of centralization and specialization, in reinforcing an already rigid school organization and standardization. Individual schools will be unable to establish close relationships with their neighbor.

Now you may not subscribe to that so I'll become more specific. I believe that a lot of the discussion of evaluation that happens at present is very good in terms of logic. Unfortunately, it is quite divorced from the reality of the classroom and the way change is introduced through the classroom. I am not surprised. In the last ten to fifteen years most educational innovations that have been introduced from the top have failed. They have failed because they have not been able to deal with the complexity of the varia es that are in operation within classrooms in individual schools.

Why do I say that? Well, first, I believe that education remains more of an art than a science. I believe that most teachers regard themselves as artists. They regard themselves in a very dynamic, a very personal interaction with a group of children. When they are working with the children they see themselves creating and moulding the same as an artist does with sculpture or with a painting. If you could accept that as an initial starting point, we get into a number of the very frightening problems in trying to deal with education as a system and introduce change on a system-wide basis.

For example, I would submit that most evaluation that is done in education is not done by any form of instrument. When a school decides to introduce this or that type of innovation they have decided before they ever introdued the innovation that the innovation will succeed. In introducing an innovation that has already been evaluated as a success, any type of information that comes out subsequently that contradicts the predetermined conclusion is doomed. "The instruments are bad, there are so many varialbes that we can't decide." Maybe the standardized test showed this or that, but the teacher will say "I saw a smile on the lips of a child who's never paid any attention in class to this time." Or there's even the light in the eyes. I've never been able to see the light in the eyes but I know teachers do refer to this. There have been studies done on excellent teachers and excellent teachers on the whole have reported that they make little if any use of standardized tests. There are also many reasons why it has been difficult to replicate, for example, demonstration projects. "There are not enough dollars, that school has more staff than us, they have carpeted corridors, they have no walls and we do have walls." The superintendent is for or against, the trustees are up or down. Therefore, whatever the minimal level of information based on a demonstration project, it is very difficult for another schools' staff to use it in whole or in part.

I would submit that educational research, and I use the term research in a very loose sense, has been extremely limited despite years, decades, hundreds and thousands of hours in various courses on tests and measurements with which teachers have been inflicted. Very few teachers are interested in devising tests and I would say even fewer are interested in looking at the results of standardized tests. And so after years of teachers' making little use of standardized testing maybe we should be looking back - Why do teachers reject the taking of very general goals, pulverizing them, isolating some obscure skills and then aggregating them? Perhaps teachers refuse to admit, refuse to accept that this aggregation is the product of an educational activity for a year.

We could go further in looking at how curriculum specialists have developed skills lists with sequential layers of skills that are somewhat fallacious scientifically but we won't. I belive that there is very, very limited value in educational research. I'm not going to deal with the fact that most educational studies suggest that there are no significant differences, no statistical differences, and therefore further research is required, which begins to drive those who are in the field of action up the wall. But I do submit that the basic data base that teachers have been exposed to for some decades has had very, very little impact as far as schools are concerned.

I'll go on to deal with some of the problems that we have within Manitoba, as far as taking some of the general concerns for a better data base and placing them into action. One of the main problems that I believe we have in Manitoba among all people involved in education, be it the teacher, the Department of Education, the trustees, is a high degree of fear, prompted by a whole host of variables which are not related to evaluation. Need we mention the number of teachers in those divisions where there is a declining school population?

Whatever the case, we're dealing with a system that not only is varied on a horizontal plane, but also on a vertical plane; where we have a large number of teachers, a large number of division administrators who are operating at various stages of development. Unfortunately, most of our new schemes are directed towards those who are at the right side of the normal curve - the very energetic, the very dynamic, the very bright, the very courageous and very self-motivated, which further puts pressure on the poor fish who are operating at a minimal level, and are frightened for their skins despite the legal protections of tenure. It is this other half of the normal curve that is usually ignored, and these attempts to generalize put their insecurity into greater focus rather than building upon the strengths that they have.

Let us look at the Department of Education for a moment. Just a moment....I don't want to threaten you too much. The Department of Education has a very mixed reputation, going from somewhat helpful to, on the other hand, quite regulatory and insensitive. Even in what I call a very mild, non-threatening application of consultants, we've got a very mixed record, and on the whole not too positive in the schools of this province. What will happen if we have our evaluation experts floating out? Next, the Faculties of Education. I will throw in the cover that individual exceptions can be noted. Now....(laughter)....I would submit to you that if the faculties of Education undertook the matter of evaluation as a major activity, and if they succeed as well as they have done in the matter of preparing teachers to cope within the classroom, God help us.

School division staff? Most school divisions in this province, the administrators, the trustees of them all, are overworked. Unless there are substantial human resources devoted to them and their relief, not much can be expected at the school division level in terms of any very serious looking at alternatives, various procedures, feed-back mechanisms and so on. In this time (if you will excuse the expresseion) of Trudeau restraints, the injection of large amounts of extra resources is very unlikely.

Another group that is ignored is the lay public, although they are continuously evaluating the school informally through a variety of devices - visits to the school, coffee parties and so on. They have played an insignificant role and I believe as far as decision-making is concerned that they will continue to play an insignificant role.

So in trying to come to a conclusion on our provincial evaluation system I've come to the conclusion that evaluation should be kept modest, should be kept moderate, should be kept somewhat gross and not too fine, in order at least to provide a very minimal, minimal amount of information for decision-making by a variety of actors within the provincial system. To go beyond that, I believe, would only reinforce the trend that would make our system, which I consider not flexible enough now, even more inflexible. I point out again that evaluation has to be regarded as extremely political. Until I can see some very dramatic changes within the political environment, the way decisions are made within our educational system, I see not much hope for evaluation on a more serious note within this province.

Let me look at the hope that I have and say that I think that all educators have to be idealists. If we are not idealists, we should get out. But as I tried to write this I began to wonder if I should because I see a tremendous quantum leap between the what-is and the what-ought-to-be. I believe that our Manitoba educational system is no different from all the other systems in North America. It is a system that is marked by hierarchy - the hierarchy which places the teacher and the kid and the parent at the bottom and then winds its way up. It is a system that is marked by specialization. So while many of our new whatever-it-is's in education are always designed to support the classroom teacher, they do not. What is intended as a support and a service to the classroom teacher usually turns out to be just another instrument to define and constrain and restrict the modicum of flexibility at the classroom level. And it is only at the local level, the school, the classroom level, where the whole child is taken into consideration. It is not the mathematics child nor the social studies child nor the evaluation child, it is the whole child. But our organization, our decision-making structure works in complete opposition to that.

We're dealing, I think with a time when there is a lack of confidence expressed by a large number of people in education. We have the basic zealots, the radical reformers, the disenchanted minorities, the poor, who all are looking for a system that will become more sensitive to their needs. Now I know that this sounds rather trite, but I do believe that education is a public enterprise, and participation should be afforded so that all citizens could have the opportunity for action and initiative in this domain of life. They should have a share in the definition of the goals. Our education system, by the way it is organized, seems to deflect not only teachers but also citizens from having a more direct relationship with decision-making. Our administrative procedures, too, are somewhat oblivious to this. As a result we have to be looking for various types of structures that will promote opportunities for all interested parties-pupils, teachers, parents, and so on-to have a share. If we could develop certain types of models of consultation and participation that would permit discussions, that would permit confrontations, but at least have the decision-making taken to the action level-the classroom-it would be possible for the information and decisions that are to be made to be translated quickly. Yet I must be realistic, and in being realistic I believe that the sketch that I have outlined has very little promise in the Province at present. I believe it has very little promise elsewhere in North America. Right now I would repeat that the prevailing mood in dealing with the swing of the pendulum will be to reinforce centralism, be it at the school division or the Provincial level, to reinforce specialization by bringing in more experts, and thereby to exclude or to make even more insignificant the role of the individual classroom teacher and the neighborhood in which that teacher works.

In conclusion, I would state that in the matter of evaluation activities in the department of Education undoubtedly in the next few years we'll move into other forms than the three general areas I sketched out at the beginning of my talk. The instruments undoubtedly will be the product of the host of variables that now affect the education system. I'm afraid that even the modest amount of evaluation that I have sketched will prove inadequate and that as the pendulum swings further over we will begin in the Department to have to respond to other forms of evaluation instruments

that will likely have very little benefit for the individuals who count most in the classroom but will be of benefit to those who are interested in nice, neat controls. The quantitative approach, for example, the use of collecting numbers, that will put this school .3 of a norm grade level ahead of another will become all-important, although I maintain it would have very little beneficial impact on classroom performance. And so what I'm suggesting to you is that I've had to wrestle with a major dilemma over the last few years. For a variety of reasons a starting point has been made and a series of steps with a self-fulfilling prophecy has resulted. On the one hand we see a growing pressure to return to basics, to fundamentals. Once that could be accepted - and I believe a large number do wish to accept that - the next step is the development, the promotion, of very nice, neat numbers games, from nice little tests and what have you. That in turn leads to nice, neat information being packaged for superintendents, for trustees, for Departmental officials, who could then use the data to promote more centralization and controls. On the other hand, I find very limited support, but a very urgent need, to make the system more sensitive to the needs of a large number of pupils who are poorly served now. To do that the next step is to create a type of decision-making structure where information is promoted, where feed-back is developed at a very local level to involve a large number of actors who now have a minor role in the decision-making structure in this province.

It is clear that I see in the dilemma that I am reacting to the artistic part on the one side and the more rigorous systematic part on the other. I think you know that I am fighting what I would call a rear guard action. That does not surprise me. The record of change, of reform in opening the schools in North America has not been bright in the twentieth century. While individual examples of the types of schools that are quite open, quite sensitive, can be cited, as a system I think it has moved all too slowly. Therefore, you should look at evaluation in terms of political analysis; at least that would provide a starting point that will question a number of the assumptions that we start with as premises in our evaluation field today I would say that I see it as a large stellar brain with a huge black void and a few interesting practices flashing about within it.

Thank you.

CONCLUDING REMARKS - TOM MAGUIRE

The time has come to start in. When I was a boy, ten years ago, at the University of Illinois, I shared an office with Peter Taylor, for whom this memorial seminar has been convened. We used to sit in this room and argue and talk, and often from outside one of the adjoining offices Gene Glass who at that time could be characterized as the cherubic, acclaimed prime Doctoral student from the University of Wisconsin, would come. He used to come down and sit there and get into arguments with Peter Taylor. He would reach the point where he was ready to leap in for the kill and suddenly - and those of you who knew Peter will recognize this - Peter would slide away and Gene would go back into the office shaking his head wondering what had happened.

That's somehow the way I feel today. I came here ready to leap in, drive the harpoon home into evaluation, and somehow it slipped away from me again. I'm left wondering why, so I'd like to entitle this concluding set of remarks, "Always Say No When Somebody Asks for an Evaluation" or "Doing Something is Not Always Better than Doing Nothing".

After yesterday's session Bob came over to me and asked why it hadn't gone well (which was my perception too). It seems to me that part of the problem was that when we start to talk about evaluation - and this is such common sense that I'm mildly embarassed to be talking about it - each person has a different picture in their heads. I hadn't been confronted with that recently until the late session yesterday.

Maybe what is needed is for someone to sit down and take the time to develop something that could be called a taxonomy of educational evaluation tasks. I haven't been doing much reading in the area of educational evaluation in the last little while, mostly because I think there's so much shaft involved that it's hardly worthwhile. What one does is waits until one hears what the Tony Riffels and the Bob Stakes have to say, and lets them do the reading; they probably read faster than I do anyway. So this job of setting up a taxonomy of educational evaluation tasks may very well have been done. But I think what we need to do, if it hasn't, is to take the time to do it so that we can separate things rather clearly, such as program evaluation, the evaluation of school systems, because it seems to me that some kind of benefit can accure from that exercise in the same way that a benefit was gained from the taxonomy of educational objectives of many years ago. Everybody complains about it, but it's very productive and very useful when you're trying to talk about educational objectives because they have some kind of common meaning. You can say this is what I'm talking about, and in a few words everybody is in roughly the same picture. Part of our problem, I think, was that we had different pictures. The things we talk about in evaluation apply more or less to the different taxonomic levels of evaluation tasks. Perhaps when people try to enunciate general statements of evaluation the very first thing they can do is to say "Would any reasonable person say this?"

What I'm suggesting we might do is, if we have general statements about evaluation, take such statements and then apply them to the various levels of this new taxonomy to see if it makes sense to talk in a certain way when referring to a certain taxonomic level.

I was attracted by Tony's characterization of evaluation. I was so attracted to it that I'll probably make it into my own and it will be completely unrecognizable. But it seems to me that it might be productive to disregard the notion of evaluation and start to pay more attention to the role of facilitation and the interaction with data collection. Perhaps what we need are training programs for facilitators. Perhaps what we need is to spend a little more time looking at the political nature of educational evaluation or that process that will be continued and see what the implications are of these political approaches to educational evaluation. In the measurement area it seems to me that the statement has been made and the assumption seems to me to be rampant, that data collection is pretty well satisfied at this point. My feeling is that it's not. What is true is that statistics, just bald old statistics, have probably far outstripped any of the conceptual methodology of evaluation. We're in good shape statistically speaking, in terms of sampling and trying to make statements about what the state of some population is about, but we're in terrible shape methodologically, I think, and that's what fifty years of educational research has been trying to tell us. Fifty years of educational research says that short of cutting off the kid's head it really doesn't much matter what you do to him in the classroom. They learn in spite of it. Maybe what we can do is push more toward anthropological methods, journalistic methods and maybe even political-scientific methods. I don't know.

The next thing that we need to do is look again at educational research because if we are as evaluators or as people carrying on that role going to make any kind of prescriptive statements to people in education, I think we had better be standing on firmer ground. Some of the things that we have found in educational research don't seem to stand up over time. You know the work that's been done by Flanders in which he's shown that there's a relationship between indirect teaching methods and achievement and that this relationship corelation is about in the neighborhood of .4. What's kind of interesting about that data is if you go back and look at teacher inter-action with individual students and calculate the co-relation between the individual student achievement and the indirect to direct teaching ratio, you find out that that corelation washes out to a mere nothing and in some cases even becomes negative. In other words, over various classes you get a kind of relationship but within classes there is either no relationship or even a slight negative relationship. The more indirect you are, it seems the lower the achievement. Except that's not the explanatory phenomenon that's going on. What's really happening is any good teacher worth his salt spends most of his time working with the dumb kids.

What am I getting at with this kind of confusing thing? Well, I'm saying that there's not an awful lot that we know. That once we find out, for instance, that a certain curriculum or a certain procedure is not working we're in a little bit of difficulty about what to tell the people to do. Some people are going to have to sit down and spend some time thinking about that whole role of educational research and the possibilities for educational research as some kind of generalized scientific endeavour. Wahlstrom almost put his foot on it yesterday when he tried to direct our attention to method of evaluation. I'm going to leap in with both feet and say that we need some stuff that I'm calling evaluation research and that's the research on evaluators. What we need is some bright sociologist to spend a little bit of time observing evaluators as they come to conventions or as they go out to work so that we can begin to say - what is it that educational evaluators do apart from what they say they do? I think these are two quite different things. This is tied into a question that we didn't get into yesterday afternoon that I think Tony wanted to talk about and that was - how does the

context of the educational evaluation predetermine the outcome of the evaluation itself? Tied into that I think we need some kind of research on evaluators.

The final need, and I'm having trouble with this one myself, is to put to death the notion of the superevaluator or Captain Summative. My guess is that almost nothing in education is summative. In evaluation the point is this: we are much more like a canoe going down a river. It really doesn't much matter to say that that was the perfect thing or that was the terrible thing because the whole situation is constantly changing. The question that the evaluators should be addressing themselves more to is what should we do now? And why should we do it?

All this takes me to the airport with a couple of questions. It seems to me that as I leave and perhaps as you leave, we're filled with feelings of confusion, maybe disillusionment, certainly a lot of questions. I think it's fitting this conference should be the Peter Taylor Memorial Conference because in the eight or so years I knew Peter, I think that I could characterize half our encounters as finishing with that same kind of sense of confusion, occasional bewilderment, but very often a sense of questioning, a sense of exhilaration, at having spent the time threshing away at this.

With that thought I'd like to thank you for inviting me. I hope that for these questions you'll be providing answers to me, and I'll grade them and get them back to you by the end of the week.